WILL RAISING THE BAR RESULT IN GREATER READING GROWTH

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF EDUCATION

BY

STACY LYNN SMITH

DR. MARILYNN MARKS QUICK – ADVISOR

BALL STATE UNIVERSITY

MUNCIE, INDIANA

MAY 2019

SIGNATURE PAGE

WILL RAISING THE BAR RESULT IN GREATER READING GROWTH

A DISSERATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF EDUCATION

BY

STACY LYNN SMITH

DISSERTATION ADVISOR: DR. MARILYNN MARKS QUICK

APPROVED BY:

_____	_____
Committee Chairperson	Date
_____	_____
Committee Member	Date
_____	_____
Committee Member	Date
_____	_____
Committee Member	Date
_____	_____
Dean of the Graduate School	Date

BALL STATE UNIVERSITY
MUNCIE, INDIANA
MAY 2019

ACKNOWLEDGMENTS

By far, the person who guided and supported me throughout this entire journey has been Dr. Quick, my dissertation chair. Thank you for not giving up on me, pushing me, providing me with constructive feedback, and encouraging me through these last several years. You are far more than just one of my professors, you have become a mentor, friend, and counselor. I have gone through many significant life changes throughout this journey and would not have completed this work without your unending support and guidance.

Dr. Boyland, Dr. Jackson, Dr. McKinney, Dr. Eouanzoui, and Dr. Jones, thank you for your time, expertise, patience, and support throughout this process. I am very appreciative for your guidance, input, and contributions to improve my work.

My parents, Larry and Ginny, have modeled for me perseverance, hard work, and integrity throughout my entire life. Mom and dad, without your selfless and unwavering support, I would not have completed this journey. I hope that you know that your belief in me and unending love made this possible.

My children, Tori and Josh, can't remember a time that their mom wasn't both working full time and pursuing advanced degree work. I am hopeful that the accomplishment of this long and difficult journey will bless you in years to come, and that you will know that all things are possible.

Finally, to my loving husband Jay, who entered into this process after it had already been put into motion. Thank you for your complete and total support over the last several years. Thank you for not letting me give up and encouraging me through all of the ups and downs. I love you completely and cannot fully express how much I appreciate all of the many ways that you have sacrificed to help ensure that this came to fruition.

DEDICATION

To my husband, Jay, whose unending support, love, and encouragement were crucial to the completion of this work.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

The placement of students in small guided reading groups at their instructional reading level is universally accepted as best practice by the majority of educators, reading specialists, and interventionists (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996). For decades, Emmett Albert Betts (1946) was considered the forefather of how to determine a student's independent, instructional, and frustration reading levels through his work on informal reading inventories, or IRIs (Boley & Pennock 1975; Ekwall, 1976; Johns & Magliari, 1989; Pikulski, 1990; Powell & Dunkeld 1971; Williams, 1959). However, there are some scholars questioning the origin of the research data used regarding the percentages recommended for the reading levels set forth in Betts' 1946 book, *Foundations of Reading Instruction* (Cooper, 1952; Johns & Magliari, 1989; Powell & Dunkeld, 1971). More recently, according to the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), the level of text complexity must be raised for all grade levels in order for the United States to compete internationally. However, would raising this bar of text difficulty close the gaps in reading achievement on its own or do we also need to raise our expectations of what constitutes an instructional reading level? Is the percentage range for which we base an instructional reading level accurate and rigorous enough? Will students become frustrated and achieve at lower levels if the text is too difficult for them? These questions are the very heart of my proposed study and have driven me to continue to ask questions and dig deeper.

Purpose of the Study

One purpose of this study was to determine which instructional reading level text (one that matches, is below, or is above a student's found instructional reading level) used in small group guide reading fostered greater reading achievement and growth as measured by NWEA. The term *match* in this study refers to when the difficulty level of the text to be read is within the instructional reading level range of a student. A subsequent purpose of this study was to identify the relationship between whether a student's found instructional leading level was within, above, or below the typical grade level range and their reading achievement scores and growth. The independent variables in this study were gender, race, SES, Limited English Proficiency, special education, the found instructional reading level of students, whether the found instructional reading level was within, above, or below the typical grade level range, and the instructional reading level of texts used with students. The dependent variables were the reading achievement of the students as measured by NWEA RIT and nationally normed percentile on the pre and post-tests and the achievement growth between the pre and post-tests.

Another purpose of this study was to determine the perceptions of the teachers and the school leaders as the teachers implemented more difficult texts with students during small group guided reading. In addition, the perspectives of the principal and professional liaison were sought regarding how they might best support literacy initiatives and how the results of this study might be expanded to other settings.

Research Questions

The following research questions were addressed in this study:

1. What is the relationship between the instructional reading levels of the texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading growth as measured by NWEA?
2. What is the relationship between the instructional reading level of texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading achievement as measured by NWEA?
3. What is the relationship between whether students' found instructional reading levels are within, below, or above the typical grade level range and their end of semester reading achievement and growth on NWEA?
4. What are the perceptions of the teachers as they increase the instructional reading level of texts used with some students?
5. What are the perceptions of the principal and the professional development liaison in expanding the results of this study to other teachers?

Significance

The emerging work regarding the rigor of text and instructional reading level in the areas of dyad reading and fluency (oral) reading have had many results that support my question regarding whether or not our current understanding of instructional reading level is rigorous and extensive enough. In addition, when considering the peer model for dyad reading, coupled with the reading gains achieved using texts at higher instructional reading levels in the dyad study

(Morgan, Wilcox, & Eldredge, 2000), the impact of utilizing texts at higher instructional reading levels within teacher-led guided reading is something to consider researching.

Furthermore, Betts' (1946) suggestions of the percentages to use when determining a student's instructional reading level have served as the baseline, without much question, of what educators have used to determine instructional reading levels for more than 50 years. There is evidence to support the questioning of the origin of Betts' instructional reading level percentages. It appears that we may need more objective investigations to be conducted in this area. The ambiguity surrounding the origin of the percentages used to determine instructional reading levels and the dyad study that indicated greater reading achievement may be reached by using texts at a higher instructional reading level have both influenced the research questions brought forth in this study. I am curious as to whether or not the traditional percentage that describes instructional reading level is rigorous enough; more specifically, I am wondering if teacher-led guided reading groups will achieve higher gains if instructed with a text higher than practitioners' current understanding of instructional reading level.

If the analysis of my proposed study supports the dyad reading findings (Morgan et al., 2000), which indicated that utilizing a text at a higher instructional level increases student achievement and growth, that would provide increased awareness that teacher training and practice may need to be altered. However, if my proposed research doesn't support the dyad study (Morgan et al., 2000) then my research may support the current practice that instructing students with texts at their instructional reading level is appropriate to make the most gains and growth in reading achievement.

Delimitations

This study included second grade classrooms in one school district. The research that was previously done by Morgan, Wilcox, & Eldredge (2000) in the area of dyad reading and text levels were conducted within second grade classrooms, which is why I chose to narrow my study to this grade level.

In addition, the location of the study was another delimitation. I chose to conduct my study in the mid-western part of the United States, which enabled me to more easily conduct my one initial observation and the multiple interviews with teachers and school leaders throughout the life of the study within a 60-mile radius of my school setting.

Lastly, I chose the setting of small group guided reading as the component of the school day in which to track and monitor the impact that different levels of text had on student reading achievement and growth. I wanted the environment to be one in which the teacher could maximize individualized instruction in order to more easily provide scaffolds and supports. I could have chosen the setting to occur during mini-lessons or fluency groups; however, these didn't appear to be as appropriate when considering the most appropriate time during the literacy block for a teacher to most easily match a reader to his or her instructional reading level and effectively monitor and support his or her learning.

Definitions

Independent reading level: the level at which a student can read and comprehend a text independently and with ease. The accuracy rate is 95-100% with excellent comprehension (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996).

***Instructional reading level*:** the level at which the student needs support and scaffolding from a teacher. This is considered the ideal level of text for reading growth to occur. The accuracy rate is 90-94% with satisfactory to excellent comprehension (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996). In this study this term is abbreviated to IRL.

***Frustration reading level*:** the level at which a student experiences frustration and the text is considered too difficult for reading growth to occur. The accuracy rate is below 90% (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996).

***Informational reading inventories (IRIs)*:** a diagnostic assessment given to help to determine a student's instructional reading level, as well as comprehension, word recognition, and word meaning (Pikulski, 1974).

***Running record*:** a formative assessment, similar to an IRI, used by the teacher to calculate a student's independent, instructional, and frustration reading level by taking oral reading and its rate and accuracy into account (Clay, 2002).

***Guided Reading*:** small group reading instruction led by a teacher which consists of students reading the same level of instructional text and displaying similar reading behaviors. This structure is meant to differentiate reading support in order to further develop reading proficiency (Fountas & Pinnell, 1996; Rubin, 2011).

***Readability Measures*:** formulas used to assess the difficulty of the readability of texts, which gives a measure of text complexity, which is then then used in order to correlate to a student's reading level (Benjamin, 2012; Fisher, Frey, & Lapp, 2012).

***Text Complexity*:** a measure of how difficult or easy a text is to read and comprehend. A widely used measure of text complexity is the Lexile level that can be applied to both texts and the reading level of a student (Benjamin, 2012).

***Dyad reading*:** a form of unison reading that involves two students, one as the lead reader and the other one as the assisted reader, who has weaker decoding skills. Dyad reading is also known as buddy reading. (Morgan, Wilcox, & Eldredge, 2000).

***Fluency*:** oral reading of a text measured by the by rate, accuracy, and expression used by the student (Benjamin & Schwanenflugel, 2010). Fluency is a bridge to comprehension and is an area practiced within small group guided reading.

Professional development liaison: a professor at a local university that partners with a school in order to facilitate the professional development of the teachers, as well as to place and supervise student teachers from the university into the school setting.

Scaffold: in reading instruction, a scaffold is an additional support provided to a student who is reading a text that helps him or her have success in reading the text. Some examples of scaffolds may include: allowing the student to pre-read a text, listening to the student read the text and prompting with strategies as needed, providing a video to watch on the subject presented in the text, pre-teaching vocabulary, or even helping a student complete a graphic organizer, or writing activity that is connected to text.

Organization of Study

The remainder of this study is organized into four additional chapters, a bibliography, and appendixes. Chapter 2 provides a review of the literature regarding the history of instructional

reading levels and current literacy practices surrounding the use of instructional reading levels and the teaching of reading. Chapter 3 explains the research methods and design of the study, including the instruments used to conduct the study, the sampling procedures, and data collection process. A discussion of the findings and the analysis of the data are found in Chapter 4. The recommendations, conclusions, and summary of the study are located in Chapter 5. The bibliography and appendixes conclude the study.

CHAPTER 2: LITERATURE REVIEW

This chapter provides a review of the literature pertaining to instructional reading levels and their use within small group guided reading. It begins with the history of how reading levels where ascertained through the work and writings of Betts (1946). The next four sections discuss additional reading level assessments, text readability measures, and text complexity, including both current methods and implications regarding these topics. Furthermore, emerging approaches in literacy instruction in the areas of dyad reading, guided reading, and fluency are discussed. Finally, a modest portion on reader interest and teacher effectiveness is discussed, along with the role that leadership plays regarding effective literacy change in the classroom. The chapter ends with conclusions and a summary.

History of Reading Levels

The review I have conducted of the literature available regarding the historical context of the independent, instructional, and frustration reading levels has been thought-provoking. I was surprised to find that our current system of using instructional reading level percentages to match children to texts in order to maximize their reading potential or close their achievement gap, is mostly based on the work of educational researcher Emmett Betts's 1946 book, the *Foundations of Reading Instruction*. The origin of these three basic levels of reading and the assessments that have been used in the diagnosis of these levels are important to understand and consider.

Birth of independent, instructional, and frustration leveling system. In 1946 Betts published *Foundation of Reading Instruction*, which laid the groundwork for many crucial aspects of effective reading instruction (Betts, 1946). Among these lie systematic instruction, differentiated instruction, and diagnosing specific reading needs through the use of Informational Reading Inventories (IRIs) in order to place students in instructional reading materials for direct instruction, versus independent (basal) or frustration levels (Allington & McGill-Franzen, 2000; Betts, 1946). An IRI is an assessment given to a student to help to determine the student's instructional reading level, as well as to assess comprehension, word recognition, and word meaning (Pikulski, 1974). According to Betts (1946), after conducting an IRI, the percentages to use for distinguishing the levels are as follows: independent (basal) level >94% comprehension and >99% accurate pronunciation of words, instructional level >75% comprehension and >94% accurate pronunciation of words, and frustration <50% comprehension and inability to pronounce 10% or more of the words (p. 446-451).

The Betts book (1946) included many principles that would still be considered best practices today. In searching the text, I focused on finding the way in which the percentages of the three reading levels were determined. Regrettably, I was unable to find any research conducted by Betts included in his text to substantiate his claims regarding the origin of these percentages; however, he did mention the works of Thorndike (1934) and Killgallon (1942) in his section that summarized his suggestions for reading levels. Furthermore, in a 1989 article titled *Informal Reading Inventories: Are the Betts Criteria the Best Criteria*, authors Johns and Magliari stated that the criteria for reading levels "was based, at least in part, on a study done by Killgallon (1942) under the direction of Betts" (p. 124).

The unpublished doctoral dissertation study done by Killgallon at Pennsylvania State University in 1942 has been challenging to obtain and I was unable to locate it. In addition, Johns (1991) wrote an article based on an interview with Betts prior to his death in 1987 in which he wrote that "Betts directed Killgallon's 1942 dissertation, generally regarded as the empirical basis for the oft-quoted numerical criteria for word recognition and comprehension" (p. 493). The sample size of the Killgallon study, according to Betts (1946) and Johns & Magliari, (1989), was comprised of only 41 fourth grade students from central Pennsylvania. Johns (1991) shared that in his interview "Betts, however, said that missing one word in 20 (the instructional level) was already established by his own studies and the work of Thorndike and that Killgallon went with the criteria already established" (p. 493).

According to Powell (1971), very few investigations regarding the validity of the criteria used for determining the instructional reading level have been done. "Only three studies, Killgallon (1942), Cooper (1952), and Powell (1969) have attempted to offer data to support the criteria they suggest and these three studies conflict sharply in the criteria they recommend" (Powell, 1997, p. 638). Paolo (1977) stated that regarding the research surrounding Betts' research, "evidence was conflicting as to how the criteria were established and some investigators have questioned their general acceptance" (p. 22).

When considering the information thus far, there is always a possibility the research used to validate Betts recommendation of reading level percentages is in a work of his that I was unable to review and or locate. Betts has supporters for his work with IRIs and the recommendations he set forth for the criteria when determining instructional, frustration, and independent reading levels (Ekwall, 1976; Pikulski 1974; Williams, 1959). However, there were some who were starting to have concerns regarding the research used to determine the

percentages for determining reading levels starting back as early as the 1950s (Cooper, 1952; Powell & Dunkeld 1971; Boley & Pennock, 1975; Paolo, 1977; Johns, 1989). However, even though some are questioning Betts' criteria, "many authorities not only accept them uncritically, but also regard them as definitive of reading levels at all age and grade, and are used as the baseline" (Powell, 1971, p. 638).

In summary, Betts' (1946) suggestions of the percentages to use when determining a student's instructional reading level have served as the baseline, without much question, of what educators have used to determine instructional reading levels for more than 50 years. There is evidence to support the questioning of the origin of Betts' instructional reading level percentages. It appears that we may need more objective investigations to be conducted in this area. In addition, there is a more recent study in the area of reading, discussed further in this paper, which implies that a more rigorous instructional reading level might result in higher reading achievement. The ambiguity surrounding the origin of the percentages used to determine instructional reading levels and a recent study that indicated greater reading achievement may be reached by using a higher instructional reading level, have both influenced the research questions brought forth in this study.

Reading Level Assessments

In addition to the IRIs that Betts (1946) used and promoted in his book, another well respected informal reading assessment used to find a student's instructional reading level was developed by Marie Clay in the 1970s (Dunn, 2010). This program was called Reading Recovery (RR) and was an intervention program focused on first grade readers (Dunn, 2010). However, from RR came the assessment that is now used in many classrooms in order to ascertain a

student's range of reading levels, the running record (Clay, 2002; Dunn, 2010). A running record is used by the teacher to calculate a student's independent, instructional, and frustration reading level by taking oral reading and its rate and accuracy into account. Our current application of percentages in order to assess independent, instructional, and frustration or hard reading levels have been adopted by the majority of reading specialists, teachers, and are reflected in curricular resources. The common percentages used for word accuracy are: 95-100% independent, 90-94% instructional, and below 90% frustration or hard (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996). The common practice in the majority classrooms is to instruct students at the oral reading accuracy rate of 90-94%, which is considered instructional reading level (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996). An example what this might look like is provided in Figure 1.

Oral Reading Accuracy	**Reading Level**
95-100%	Text is at student's **independent** reading level
90-94%	Text is at student's **instructional** reading level
Below 90%	Text is at student's **frustration** reading level

Figure 1. Reading accuracy percentages and student reading level.

In comparison, an IRI, as mentioned above in Betts' work, is a more extensive informal assessment because it also contains comprehension questions in its formula for calculating reading level, as well as student interests. However, both running records and IRIs are still the most widely forms of informal reading assessments used to calculate the reading level of students. According to Ross (2004), "effective school studies consistently report positive

associations between high student literacy and engagement in systematic classroom assessments procedures, particularly running records" (p.191).

Interestingly, Ross' (2004) Canadian controlled experimental study, consisting of 75 K-6/8 schools comprised of 2,800 third grade students, confirmed that schools implementing running records to diagnose student needs and monitor progress outperformed similar schools in the same district in both reading and writing achievement. Half of the schools were assigned to the running records treatment and the other half to action research, which was comprised of leadership deciding on literacy initiatives to focus on for planning. Ross (2004) controlled for prior school achievement and collective teacher efficacy. "The running records intervention accounted for 12% of the between-school variance in reading and 7% in writing, confirming the correlational findings of effective schools' research" (Ross, 2004, p. 186).

A strong body of research indicates a link between the assessment of reading level and effective instruction, along with a myriad of other teacher behaviors. This inference is supported in the writings of Calkins (2001), Clay (2002), and Fountas & Pinnell (1996), which all encourage the practice of conducting running records as part of an effective literacy program. In addition, the acknowledgment that effective literacy instruction is made up of a complex interaction of components has a broad range of supporters within the field of literacy (Pressley, Wharton-McDonald, Allington, Block, Morrow, Tracey, Baker, Brooks, Cronin, Nelson, & Woo, 2001). Therefore, in order to help teachers to become more effective, a focus on engagement and authentic reading and writing, coupled with the systematic process of assessing reading levels and differentiating instruction accordingly, would be considered best practice (Pressley et al., 2001)

Calkins (2001) suggested that the identification of a student's reading level relies on everyday observations, running records, informal reading inventories (IRIs), and miscue analysis. These all have a place in the ebb and flow of a student's growth in reading. However, in practice, many teachers may rely on too few data points when making decisions about the level of texts to use when instructing students. According to Rubin (2011), "research has questioned the wisdom of using results from only one tool to pass judgment on how well students comprehend text" (p. 606). Therefore, Rubin (2011) recommended that teachers compare the scores from multiple assessments, such as running records and IRIs in conjunction with standardized assessments in order to make the most accurate decision on the instructional reading levels of students.

This thought was reiterated by Ford and Opitz (2008) when they wrote, "small group reading instructional of the past often relied solely on end-of-level assessments built into programs, but teachers implementing guided reading need to become increasingly more expert at continuously conducting assessments and interpreting results" (p. 322). They went on to encourage the use of a variety of assessment techniques when trying to ascertain a student's level of reading, honoring the complexity of the task.

The lack of evidence surrounding the research regarding the origin of accuracy percentages used to ascertain an independent, instructional, and frustration reading level gives cause to my question regarding whether or not our current understanding of an instructional reading level is accurate and rigorous enough to maximize the reading growth of students. For example, Powell (1971) recommended in his study, *Validity of the IRI Reading* Levels, that an accuracy rate of closer to 85% may be a better match for increasing reading achievement, which is a direct challenge to the current understanding of 90-94% as an instructional reading level

range. In addition, the use of running records and IRIs to match students to instructional texts for differentiated instruction in guided reading may need to be woven together with other data points in order for teachers to effectively group students and match them to texts. Furthermore, if the current range of accuracy used to determine an instructional reading level appears to be either not rigorous or extensive enough, these reading assessments may need to be adjusted accordingly and more complex texts may need to be used with students.

Text Readability Measures

The practice of using a type of leveled text with students dates back to McGuffey's Readers, which are believed to be the first set of leveled basal readers used for reading instruction in America from the late 19th century to the early 20th century (Betts, 1946; Kontovourki, 2012; Pitcher & Fang, 2007). McGuffey's Readers were basal readers that were written to project graduated levels of difficulty with each progression in school, not individual reading levels. However, from the early 1920s until now researchers have created hundreds of readability formulas in order to match text difficulty or complexity, sometimes referred to as text readability, to the reading levels of students (Hiebert, 2011; Benjamin, 2012). The vast majority of these formulas include a combination of syntax, the number of words per sentence, and semantics, the number of syllables per word or word familiarity (Hiebert, 2011; Benjamin 2012).

One of the most popular text readability measures used today is the Lexile framework. It was created in the late 1980s, but is still considered an effective measure of text complexity (Benjamin, 2012). According to Benjamin (2012), the Lexile scale was created to complement our understanding of reading comprehension, created with a high degree of construct and predictive validity. The Lexile scale takes into account word frequency, which is semantic, and

sentence length, which is a way to measure syntactic complexity. However, the main difference between other quantitative readability measures and the Lexile system is that it includes matching a reader with the best level of text. According to Benjamin (2012),

> The appeal of the scale for wide use seems to be largely based on its application: a person receives a Lexile score based on his or her ability to answer comprehension questions correctly; a text also receives a Lexile score. If the person and the text are matched, then the person has a 75% chance of answering a comprehension item correctly for the text. Then the teacher can look at the Lexile score for a text and determine whether or not that text would be appropriate for a student based on the student's Lexile score. (p. 67)

Current practices and issues with text complexity measures. The consideration of text complexity measures are of great interest currently due to the recent adoption of the Common Core State Standards (CCSS) by the majority of the states in the United States. The resurgence of interest in the materials students read have been considered before; however, this renewed focus on text complexity is compelling us to go deeper and consider multiple aspects, not just text difficulty (Fisher, Frey, & Lapp, 2012). The CCSS has made text complexity one of the foundational concepts regarding student achievement in reading and has heightened the importance of the consideration of the materials put in the hands of students (Hiebert, 2011). The Common Core State Standards document states that "while reading demands in college, workforce training program, and life in general have held steady or increased over the last half century, K-12 texts have actually declined in sophistication" (NGA Center & CCSSO, 2010, Appendix A, p. 2).

For these reasons, the CCSS has a three-part model for measuring the complexity of texts. One piece of the model is called qualitative because it is concerned with aspects such as meaning, demands, and structure. A second part of the model is quantitative, which is measured using the Lexile format. The third part of the model is called reading and task, which takes the professional judgment of the teacher to evaluate (NGA Center & CCSSO, 2010).

However, according to Hiebert (2011), the three-part method of determining text complexity makes sense, but all the components had not been fully vetted prior to the release of the ELA CCSS. Therefore, how practical and easy would it be for teachers to implement this three-part system into their practice? In addition, how systemic would this process be across the nation? The CCSS Appendix A (2010) appears to mostly use Lexile levels for calculating text complexity. The Lexile bands have been reconfigured in order to establish a progressive stair-step gradient that extends from beginning reading to the college and career level (Hiebert, 2011). The CCSS has taken these Lexile levels and created bands for age ranges and put them in order from beginning reader to a student who is exiting the K-12 school system into college or a career.

Implications for reading level placement/assessment of text complexity. A study conducted by Pitcher and Fang (2007) revealed, through an analysis of sample leveled texts from the Wright Group, that the leveling system used in this series is not a very reliable indicator of text difficulty and that the quality of the texts is not consistent throughout the levels. Pitcher and Fang (2007) stated that "no method of estimating text difficulty, either readability formula or leveling systems, has been found to be perfect so far; it is ill-advised to rely on such a measure in determining text difficulty and hence reader-text match" (p. 50). The CCSS's recommendation of the three-part process for deciding text complexity seems to validate this statement from

Pitcher and Fang, since it includes the importance of professional judgment through the reader and task component.

In their book *Text Complexity: Raising Rigor in Reading*, Fisher, Frey, and Lapp (2012) also agreed that in order to have the best match between readers and texts, the quantitative, qualitative, and reader and task considerations must all be considered. The interaction between the reader's skills and the text is important to understand as this helps to determine the readability of the material. Therefore, a teacher's focus on matching readers to texts is highly important. Fisher, Frey, & Lapp (2012), stated, "We suggest that more difficult texts with scaffolded instruction should become part of the classroom equation. To ensure that students read complex texts, teachers have to revisit how they match readers with texts and tasks" (p.5).

Because of the findings of both Pitcher and Fang (2007) and Fisher, Frey, & Lapp (2012), teachers may want to consider the recommended level of text for a child to be a starting point, not a definitive match. The determination of text complexity tied to the instructional needs of an individual reader is extremely complex. It is more than just semantics and syntax; "comprehensibility resides in the interaction between reader and text" (Pitcher & Fang, 2007, p. 50). This means that other than the quantitative aspects of text, a teacher needs to consider the reader's age and interests, along with task and purpose. Therefore, the more thoroughly teachers know the interests and backgrounds of their students the more successful they will be when making decisions regarding which texts to match with which students (Pitcher and Fang, 2007). Once again, the writers of the CCSS seem to agree with the views of Pitcher and Fang and Fisher, Frey, and Lapp (2012) since they are all recommending a three-part approach for

determining text complexity with the common goal of students achieving deep meaning and understanding of text.

In summary, the data shows that the difficulty level of text impacts the act of reading, both the level of engagement and the activities embedded within it (Donne, 2011; Fisher, Frey, & Lapp, 2012). Also, Shanahan and his colleagues (2010) reported in their *Improving Reading Comprehension in Kindergarten Through 3rd grade: A Practice Guide* that text quality includes the level of difficulty and alignment to both student assessments and interest, and that this attention to text quality may impact reading comprehension. Therefore, the consideration of how text complexity is assessed is tightly interwoven with the reading achievement that a student will or will not make when these texts are matched to their instructional reading level. According to Fisher, Frey, and Lapp (2012) if a student is continually reading texts that are below his or her current grade level, then text difficulty is being decreased for that student over time, and he or she will not be able to access texts necessary at higher grade levels. How will this student be prepared for the demands of a career or college? The methods for matching readers and texts need to be more than a quantitative formula; they should include qualitative aspects and take the reader and his or her purpose for reading into consideration. Furthermore, increasing the amount of difficult texts read by students, with teacher supports and scaffolding, is another important aspect to consider when determining which texts to put in front of students.

Emerging Approaches in Literacy Instruction

In this section I have highlighted some of the approaches in different areas of literacy instruction that may be correlated to my proposed area of research. I analyzed emerging work in the areas of fluency, dyad (partner) reading, guided reading, teacher effectiveness, and student

motivation. However, the studies by Eldredge & Quinn (1988); Morgan, Wilcox, & Eldredge (2000) initially provoked my interest in wanting to research the concept of raising the bar of the instructional reading level in guided reading instruction. Both studies showed a plausible link between the increase of text difficulty and greater reading achievement in the area of dyad reading, sometimes considered partner or buddy reading. In both studies dyad reading consisted of the pairing of a stronger reader with a struggling reader. The students read the text orally together, with the stronger reader setting the pace and the struggling reader tracking the text with a finger (Eldredge & Quinn, 1988; Morgan, Wilcox, & Eldredge, 2000). The results and methods of these two seminal pieces are discussed in more detail in the next section.

Dyad reading findings. Dyad reading is a form of unison reading that involves two students, one as the leader reader and the other one as the assisted reader, who has weaker decoding skills. (Morgan, et al., 2000). Eldredge & Quinn (1988) studied the effects of dyad reading, or partner/buddy reading, with struggling second grader readers in order to raise achievement gains in reading. A group of 61 struggling or poor readers from five different schools were chosen for this study. The group of students was divided into either the experimental or control group. The experimental groups, or classrooms, used the dyad reading method with the classroom basal and other content area books; which Eldredge & Quinn (1988) described as the frustration reading level for the students in the experimental study who were paired with a different capable reader every week. However, the control group classrooms did not use dyad reading with their poor readers, but continued to use instructional level materials with these students and the classroom was more traditional in its basal instruction. This study lasted the entire school year. The results of the study indicated that greater achievement gains in

all reading outcomes resulted from those students involved in the dyad reading experiment, compared to the students in the matched control group (Eldredge & Quinn, 1988). The findings of the study revealed that at the end of the experiment, 84% of the dyad reading experiment students scored on or above grade level on the Gates MacGinitie Reading Test; however, only 19% of the control group reached these levels of achievement (Eldredge & Quinn, 1988).

In Morgan, Wilcox, and Eldredge's (2000) study, 51 second-grade students shown to read below the 2nd grade level were chosen for their study. The study was conducted to research at what text difficulty level poor readers can make the greatest gains in dyad reading and whether or not there is a point at which the level is too difficult even with the help of a partner. Therefore, the students in the study were divided into three groups; group 1 participated in dyad reading at their instructional reading level, group 2 participated in dyad reading two grade levels above their instructional reading level, and group 3 participated in dyad reading four grade levels above their instructional reading level. After the study was concluded, the findings showed that all three groups improved their reading ability; however, group 2 made the most significant gains in reading level, with a mean gain of 2.73. Group 1's mean gain was 1.55 and group 3's mean gain was 2.06. Since group 2 conducted dyad reading two grade levels above their instructional reading level and made the most gains, it begs the question regarding how this type of situation might apply within small group guided reading.

Since poor readers were shown to make the most gains when having the support of a buddy reader in dyad reading that was two grade levels above the poor readers' instructional reading levels, how might our guided reading groups, with the guidance of a teacher, using materials two grade levels above the instructional reading level, give us similar results? In

addition, the study by Morgan, Wilcox and Eldredge (2000) showed that the group with the smallest amount of gains was the group of dyad readers that used the instructional reading level texts. What would have happened if a group of dyad readers used texts three grade levels above instructional? How much more gain, if any, would this group have made, or would the results start to decline due to possible frustration? In contrast, was the level of text the main cause of this success, was it the dyad reading format, or was it the combination of both? When considering the structure of small group guided reading, are we truly pushing our students enough within a teacher-supported structure such as this, in both the diagnosis of instructional reading level and therefore the level of text and reading tasks?

Guided reading. In contradiction to the studies above, some researchers have concluded that the instructional reading level is the optimum level of instruction for student growth (Gickling & Armstrong, 1978; Treptow, Burns, & McComas, 2007). In conjunction, the 2008 National Survey of Guided Reading Practices (Ford and Opitz, 2008) promoted the use of instructional level reading materials as being level of text to foster the most growth in reading achievement, which is also reiterated in the work of Fountas and Pinnell (1996).

So what is guided reading? According to Fountas and Pinnell (1996) guided reading "is a context in which a teacher supports each reader's development of effective strategies for processing texts at increasingly challenging levels of difficulty" (p. 2) with the ultimate goal being for each child to use reading strategies independently. Small group guided reading typically consists of four to six students who have a similar need in reading, most commonly by instructional reading level. The level of text encouraged for use in guided reading by Fountas and Pinnell is text that is matched to the instructional reading level of each student. Rubin

(2011) agreed that teachers who utilize instructional-level reading materials within guided reading, and offer individualized support to students in this context, maximize their potential to learn. Furthermore, there are many authors whose common practice includes forming guided reading groups based on the instructional reading level of the students (Fountas & Pinnell, 1996, Ford & Opitz, 2008).

Guided reading groups are rooted in the concept that student groups would be flexibly grouped and that membership would not be static (Fountas & Pinnell, 1996; Ford & Opitz, 2008). However, the *National Survey of Guided Reading Practices* revealed that the survey responses seemed to show that we have not yet successfully mastered the implementation of dynamic student groups within small group guided reading (Ford & Opitz, 2008). Furthermore, Ford and Opitz (2008) went on to say that "they (teachers) must be careful not to return to what got us in trouble in the past—the inflexible use of homogenous small groups" (p. 316). A reconsideration of how students are placed into groups may be necessary. Even though students are at a similar reading level, their needs are almost always different. It may be more effective to group children of different reading levels and base small group instruction on like needs or goals in reading (Boushey & Moser, 2009; Ford & Opitz, 2008). This other way of grouping students may help to avoid the static membership and homogenous grouping that were revealed to be common practice in the 2008 National Survey by Ford and Opitz.

Fluency (oral reading) findings. Hiebert's (2005) fluency study was conducted with second grade students. Fluency is the oral reading of a text measured by the by rate, accuracy, and expression used by the student (Benjamin & Schwanenflugel, 2010). The students were divided into three groups: a control group that continued to teach the district's literacy program,

an experimental group that incorporated repeated readings using literature, and a second experimental group that infused repeated readings using content, or informational, text. The results of the study revealed that the group of second graders who used informational text in the form of the 2nd grade science and social studies texts for repeated readings, outperformed both the control group and the other experimental group. An interesting component to this study was that the group who conducted repeated readings of the informational text outperformed the other two groups in about half of the amount of time. The possible impact of using more rigorous instructional level texts with students in a guided reading setting, since informational text is typically at a higher difficult rate regarding text readability, may be something to consider. In conjunction, the recent research of Benjamin and Schwanenflugel (2010) regarding text complexity and oral reading prosody, or expressiveness, showed that the reading of more complex text enhances the connection between fluency, prosody, and comprehension. They concluded that "measuring reading prosody from a difficult text will serve as a better indicator of a child's general fluency than measuring it from an easy text" (Benjamin & Schwanenflugel, 2010, p. 400).

Additional issues for consideration. In his 1934 article, "Improving the Ability to Read," Thorndike suggested that the methods we choose to use to improve reading ability should also increase a student's interest in reading. This historical link between reading achievement and student motivation, or interest, is of high importance and still considered an important aspect of reading instruction (Allington & Gabriel, 2012; Boushey & Moser, 2009; Ford and Opitz, 2008; Gambrell, 2011; Moley, Bandre, & George, 2011; Kennedy, 2010). The renewed focus on text complexity through the CCSS honors this research. The interaction between the reader and

text might be one of the most important elements to consider. According to Fisher, Frey, and Lapp (2012), "the most important factor, the reader, is what makes the text come to life" (p.77). When a student is intrinsically motivated to read a text, they find a way to accomplish it, exceeding our expectations of what they are capable of comprehending. In a 1997 study, Wigfield and Guthrie found a 300%-time difference spent on reading between 4th and 5th grade students who were motivated to read as compared to students who were unmotivated to read. Therefore, the importance of student interest and motivation regarding text selection is another aspect to consider. Whenever possible, teachers should place complex and interesting texts in the hands of students in order to increase reading engagement, which should in-turn have a greater impact on a student's reading achievement.

In addition to student motivation and interest, the impact of an effective teacher in the classroom shouldn't be ignored. Recent research has shown that an effective teacher has a high impact on student achievement (Marzano, Pickering, & Pollock, 2001; Reeves, 2010). Effective teachers provide the needed scaffolds and instructional supports for students when instructing reading, which allows them to be more successful reading and accessing texts that may have initially been identified as a frustration level or hard. (Fisher, Frey, & Lapp 2012). The ability of the teacher to know how much support to provide in order to develop increased student independence in reading more difficult texts is a skill that effective teachers develop. Fisher, Frey, & Lapp (2012) also shared that a student shouldn't be limited to either a low-level text to read on his or her own or to struggle with a difficult text without teacher support; however, they suggested, "as students progress, they should be given increasingly challenging materials and taught, encouraged, and supported to use deeper skills of analysis" (p.8). This happens when an

effective teacher provides the scaffolds and supports for students while they are engaging with texts. Therefore, the importance having an effective teacher in the classroom supporting students during reading instruction is another crucial area for consideration.

Leadership for Effective Literacy Change

Effective leadership is highly important regarding the degree to which any instructional practice, such as a change in literacy instruction, is adopted and integrated into actual teaching, which in turn impacts student learning. Guskey's (2002) work on *change process* indicates that change in practice begins with professional development, which leads to a change of a teacher practice in the classroom, followed by a change in student learning outcomes, which then finally ends in a change in the attitudes and beliefs of the teachers. It is the change in attitudes and beliefs that helps to solidify change in a practice.

The role of the principal as the instructional leader of the school is one that has also proven to have a very high impact on student learning, with an overall effect size of 0.42 in the meta-analysis study conducted by Robinson, Lloyd, & Rowe (2008). The effect size in this study was a measure of the effectiveness of the role of the principal as instructional leader and the impact this role had on student achievement. Their study indicated that there was an important difference in the achievement of students when the principal acted as an instructional leader. An instructional leader is defined as a leader who is mostly concerned about maximizing the impact that all staff have on student learning, is visible in the classrooms, provides formative evaluation to teachers, learns alongside teachers, sets and maintains high expectations for both teachers and students, and is focused on interpreting data in order to gauge the quality and type of learning occurring in the building (Fullan,2014; Hattie, 2012; Reeves, 2009; & Marzano, Waters, & McNulty, 2005).

Marzano, Waters, & McNulty (2005) shared in their book on effective school leadership that it is highly important for the principal to be a leader in curriculum, instruction, and assessment. They also support the findings of Robinson, et.al. (2008) by suggesting there is a strong relationship between school leadership and student achievement. More specifically, the visibility of the principal in the classroom along with his or her ability to provide feedback are characteristics highly valued by teachers ((Marzano, et.al, 2009 & Strong, Richard, & Catano, 2008). This visibility may come in the format of walk-throughs and observations, both formative and summative. However, the impact of formative evaluation has more impact on student achievement and is considered to be a very important role in instructional leadership (Hattie, 2012; Reeves, 2009; & Marzano et. al., 2005). In his book, which was based on a study of over 800+ meta-analyses, Hattie stated that the influence of formative evaluation given to teachers by school leaders has a high impact on student achievement, with an effect size of 0.90.

Reeves (2009) spoke directly to the challenges and tasks associated with ability of a school leader to improve literacy instruction, and therefore impact student achievement. Reeves suggested that if educators improve literacy instruction, it will create a domino effect, resulting in an overall improvement in student learning across all subjects (2009). Therefore, it is important for school leaders to focus on the importance of consistency in reading instruction throughout the school. Reeves shared that an effective school leader paints a clear picture of what good teaching looks like and sounds like, especially when it comes to conducting guided reading and other elements of reading instruction. Reeves (2009) stated that if principals want to impact student achievement in literacy instruction, “they have a personal responsibility to understand literacy instruction, to define it for their colleagues, and to observe it on a daily basis” (p. 120).

Finally, leaders and teachers must collaborate and work alongside teachers to create a non-threatening environment in which teachers feel comfortable looking at their craft and the impact they are having on student achievement. Reeves (2009) reiterated that "expert teaching of reading is not following a script but rather is based on a bone-deep commitment to proficiency for every student" (p.120). The principal learning alongside teachers, being visible in the classroom, and creating a culture of reflection and continuous improvement, and being knowledgeable about curriculum, instruction, and assessment, will maximize the learning of teachers, which in turn will maximize the learning of students (Fullan, 2014).

Summary

Reading instruction is complex and by its very humanistic nature must be differentiated to account for the uniqueness of each student and the complex path that leads to success in reading. The myriad of components that have been linked to student achievement in reading are extensive. For the purposes of my proposed study the components considered regarding reading achievement and growth are as follows:

- student placement in small group guided reading groups through the use of multiple formative and summative assessments;
- the use of oral reading accuracy percentages and comprehension in the assignment of instructional reading levels to students;
- the level of text difficulty used in small group guided reading compared to the found instructional reading levels of students in the group;
- student motivation and interest in the reading of the texts;

- teacher effectiveness through levels of scaffolds and supports utilized when working with students in small guided reading groups; and
- the role played by school leadership in the implementation of literacy instruction.

Recent research in the areas of dyad reading (Eldredge & Quinn, 1988; Morgan, Wilcox, & Eldredge, 2000) and fluency (Hiebert, 2005) coupled with the uncertainty and difficulty locating the origin of Betts' (1946) percentages used to find the instructional reading level of students, have all given rise to my question regarding the accuracy of our current understanding of what oral reading accuracy percentages and comprehension constitute an instructional reading level. In addition, the above studies have all shown some evidence that perhaps students may learn and grow even more when they engage in reading more complex texts. Would putting more difficult texts in the hands of students, those found to be above their found instructional reading level, allow for greater reading achievement when teacher support is included? As suggested by Fisher, Frey, and Lapp (2012), "perhaps one of the mistakes in the past efforts to improve reading achievement has been the removal of the struggle" (p. 11). Allowing students to read more complex texts may help them develop the skill of perseverance and learn more about themselves as they struggle and eventually succeed through the reading of a difficult text (Fisher, Frey, & Lapp 2012). Therefore, more research regarding what percentages constitute a true instructional reading level may be needed, particularly within a teacher-led structure, such as small group guided reading, in which there is significant teacher support and scaffolding. This small group guided reading environment is an ideal place for students to practice struggling through more complex texts, since the teacher is there to provide support. Furthermore, how text complexity is determined and how texts are then matched with readers are both additional

possible topics for continued research. Future studies in these areas may lead to continued improvements in literacy instruction, which may increase reading growth and achievement of students and better prepare them for life in both college and careers paths.

CHAPTER 3: RESEARCH METHODS

This chapter describes the research methods used to determine the impact of matching a student's found instructional reading level with a level of text either below, on, or above their found instructional reading level during small group guided reading. In addition this mixed methods study also explains the procedures used to understand the perceptions of the school leaders, and the teachers as they implemented more difficult texts within small group guided reading, and consider the possible implications of this study. The chapter begins with a review of the purpose and research questions. The next sections discuss the research design of the study, the population and sample of the school community, the instrumentation that will be used, and the data collection process. Finally, the proposed way in which the data will be analyzed will be discussed, along with the limitations of the study. The chapter ends with a summary.

Purpose of Study

One purpose of this study was to determine which instructional reading level text (one that matches, is below, or is above a student's found instructional reading level) used in small group guide reading fostered greater reading achievement and growth as measured by NWEA, and to identify the relationship between whether a student's found instructional leading level was within, above, or below the typical grade level range and their reading achievement scores and growth. The independent variables in this study were gender, race, SES, Limited English Proficiency, special education, the found instructional reading level of students, whether the found instructional reading level was within, above, or below the typical grade level range, and the instructional reading level of texts used with students. The dependent variables were the

reading achievement of the students as measured by NWEA RIT and percentiles ranking on the pre and post-tests and the achievement growth between the pre and post-tests.

Another purpose of this study was to determine the perceptions of the teachers and the school leaders as the teachers implemented more difficult texts with students during small group guided reading. In addition, the perspectives of the principal and professional liaison were sought regarding how they might best support literacy initiatives and how the results of this study might be expanded to other settings.

Research Questions

The following research questions were addressed in this study:

1. What is the relationship between the instructional reading levels of the texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading growth as measured by NWEA?
2. What is the relationship between the instructional reading level of texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading achievement as measured by NWEA?
3. What is the relationship between whether students' found instructional reading levels are within, below, or above the typical grade level range and their end of semester reading achievement and growth on NWEA?

4. What are the perceptions of the teachers as they increase the instructional reading level of texts used with some students?
5. What are the perceptions of the principal and the professional development liaison in expanding the results of this study to other teachers?

Research Design

In the world of reading instruction, specifically small group guided reading, the processes of both teaching reading and students' reading behaviors are complex, messy, and anything but lock-step. Therefore, this study utilized a mixed methods approach with the hopes of honoring the multifaceted, complex process of teaching reading. Consequently, this study consisted of both quantitative and qualitative measures, which gave more breadth and depth to the research by enabling the questions of both *what* and *why* to be addressed when analyzing and explaining the data. A QUAN-QUAL model was used, in which both qualitative and quantitative data were gathered and analyzed simultaneously and given equal weighting. This created a more holistic and robust picture of the research findings (Roberts, 2010).

The quantitative data was gathered in a pretest-posttest design. This data consisted of the NWEA (Northwestern Evaluation Association) MAP reading RIT scores and percentile ratings, along with the Lexile ranges determined through the administration of the NWEA assessment. Even though my sample size was limited, reading achievement and growth were quantified precisely because of the instruments used. The pretest-posttest design not only allowed reading achievement to be compared from the beginning of the study until the end, it also permitted for the growth that students made in reading to be analyzed. The NWEA RIT scores, percentile rankings, and Lexile ranges were gathered at both the beginning and end of the study. The

achievement and growth achieved through NWEA was then be compared among the students regarding the percentages of time students read texts below, at, or above their instructional reading levels during small group guided reading. The use of NWEA ensured a more valid reliable assessment in which to measure reading achievement and growth. Information about NWEA is explained in more detail in the instrumentation section of this chapter.

The qualitative aspect of this study was extremely important, given the sample size and the importance of the teachers' perspectives while instructing students with higher text levels. This qualitative data was first gathered through conducting an initial observation of each teacher, followed by a series of three interviews completed at the beginning, middle, and end of the study. The initial observations helped to provide an understanding of the context of literacy instruction in both classrooms. The series of interviews allowed me to understand the data and results by taking into account each teacher's perspective, the amount of scaffolding they needed to provide students, and how their students responded to the higher level texts in guided reading groups. Subsequently, interviewing the principal and professional development liaison twice, once of the beginning and once at the end of the study, provided additional insights into the levels of scaffolding and supports teachers requested and used within the study, as well as how they, as school leaders, were able to provide support to them through their leadership. In addition, the interviews of the school leaders involved helped to ascertain how they might expand the results of this study to other teachers. In summation, the qualitative data collected and analyzed provided a more robust and holistic understanding of the results from the study (Roberts, 2010).

Sample Demographics

This study was conducted in a rural school district located in the Midwestern part of the United States, which also is close to a local university. The medium income of the residents of the town was about $10,000 less than that of the state average (USA Factbook, 2016). According to the Department of Education's website (2016), the ethnicity breakdown for the district in 2015-16 was about 97% White, 2% Multiracial, and 1% Hispanic. Of the approximately 780 students who attended this district, about 40% are on the Free/Reduced program, 11% are in Special Education, and less than 1% are English Language Learners.

The school in which the study took place was the sole elementary school in the district, with a population of around 410 students. At the time of this study, the school's accountability grade had risen every year since its initial grade of a D in 2010-11, to earning an A in the most recent published years of 2013-14 and 2014-15 (Indiana Department of Education, 2016). The school included Kindergarten through sixth grade students, serviced students with special needs, and provided opportunities for all students to develop additional skills through art, music, and band.

The demographics of the school compared to the district were very similar in almost all areas of enrollment. Therefore, a comparison between the school and the state's demographics was shown in Figure 2 (Indiana Department of Education, 2016). Please notice that there was a large difference in the ethnicity of the students. The school had a much higher percentage of White students and a lower percentage of minority students compared to the state. The percentage of special education students, English language learners, and free/reduced lunch

students were within a 3-5 percentage range, with the school's percentages a little lower than those of the state.

Student Demographics	School Population Percentages	State Population Percentages
Hispanic	1.2%	11.1%
Multiracial	2.2%	4.7%
White	96.4%	69.3%
Special Education	10.2%	15%
English Language Learners	<1%	4.8%
Free/Reduced Lunch	45.3%	48.3%

Figure 2. 2015-2016 school and state demographic data.

Literacy setting and participants. The study occurred within two second grade classrooms in the same elementary school. The setting of the study was the section of the literacy block dedicated to teacher- led, small group guided reading. Each classroom was comprised of about 17 students each, which totaled 34 students who participated. Both of the teachers in this study are in their second year of teaching second grade in this elementary school. One teacher had other years of experience teaching kindergarten and the other teacher was only in her second year altogether. The teacher with fewer years of experience taught the High Ability cluster students. The other teacher had a wide range of students, including two students with IEPs in her classroom. The teachers both willingly collaborated with the professional development liaison last year, focusing on small group guided reading as their primary topic of study. They additionally both excitedly agreed to participate in this study.

The sampling procedure used was a convenient sample based on the professional judgment of the teachers, which was necessary in order for the school leadership to agree to participate in the study. Therefore, this study did not direct teachers to use particular levels of texts with students or to group them in specific ways; it provided a format for tracking those teacher-made decisions in order to analyze the type of impact those decisions may have had on reading achievement and growth. Please refer to Figure 6 for an example of decision making. Finally, the length of the study spanned over an entire semester, which is about an 18-week period of time, ranging from January until approximately the third week of May.

The school leaders in this study were comprised of a building principal and a professional development liaison provided by the local university. The building principal was completing his 7th year at the school during this study. Prior to his role as principal, he taught 5th and 6th grade mathematics at the same elementary school. He did not have an assistant principal, which is one of the reasons that the school partnered with the local university in order to gain the expertise and assistance of a professional development liaison. The professional development liaison was a literacy professor at the local university, who both supported the professional development of the teachers in the school, as well placing and supporting student teachers in this school, as well as others. During the study, she was completing her 4th year in this position with the elementary school.

Within the school, there had been a recent change regarding the assessments administered and used to ascertain literacy information, such as reading achievement and growth. The recent adoption of NWEA by the district, which produces a nationally-normed reading score as well as a Lexile band for each student, became a new tool that the educators used to track student growth

and make instructional decisions within reading. NWEA took the place of other assessments, such as an instructional reading level benchmarking software program that was used in the past by the teachers to determine the instructional reading levels of students. Therefore, the teachers implemented a new running record/benchmark tool to use alongside their new NWEA assessment information. The benchmarking and running record tool that the teachers used throughout the study was the Reading A-Z benchmarking tool, which is part of the Learning A-Z products. Please see Figure 5 for the partial level correlation chart used with the Reading A-Z program. The full chart can be found in Appendix A. The teachers used this information, in addition to the information provided through the NWEA Reading Map RIT scores, to determine the instructional reading levels of their students and the corresponding instructional levels of texts to utilize with them during small group guided reading. The correlation chart shown in Figure 3 is a conversion chart used to correlate NWEA RIT scores to a Fountas & Pinnell reading level. The full chart can be found in Appendix B.

NWEA MAP for Primary Grades RIT Range	Fountas & Pinnell Reading Level
90 - 141	PC
142 - 147	RB/A
148 - 155	B
156 - 161	C/D
162 - 166	E/F
167 - 173	G/H
174 - 179	I/J
180 - 185	K
186 - 191	L
192 - 196	M
197 - 199	N
200 - 201	O
202 - 350	P

Figure 3. Chart created by NWEA to correlate a RIT band to an IRL level or text using F & P.

Instrumentation

This mixed-methods study utilized instruments that measured both the quantitative and qualitative data collected. The quantitative data came from NWEA's (Northwest Evaluation Association) MAP Reading assessment in the form of RIT scores, percentile rankings, and Lexile ranges. The NWEA Reading MAP data was used as the pre- and post-test measure in order to ensure a valid and reliable assessment in which to measure reading achievement and growth. NWEA is a not-for-profit educational service organization that created MAP (Measures of Academic Progress) and MAP for the primary grades, which is a computer based assessment that responds to students' responses in order to provide a detailed individualized report on each child. This individualized assessment supports the adjustment of instruction to what each student

is ready to learn. The Standards for Educational and Psychological Testing directed the development of the MAP NWEA assessments. MAP and MAP for primary grades are computerized assessments that adapt to the answers chosen by students and either increases or decreases the level of difficulty of the next question. This capability makes the assessment individualized and able to measure the full scope of student's capacity within a particular area, such as reading (NWEA MAP Assessments, 2016).

NWEA MAP results are based on an equal interval scale called a RIT band, which is a numerical scale that is independent of grade level, is designed to measure longitudinal growth over time, and connects the MAP and MAP for primary grades. RIT stands for Rasch unIT scale, which is the equal interval measure that is applied to both test item difficulty and student score. The reading RIT score yields an overall reading achievement score and a correlated Lexile band. The reading RIT is also broken down into subtests, that slightly differ between the MAP and MAP for primary grades. The MAP subtests are literature, informational text, and foundational skills/vocabulary. The MAP for primary is broken down into foundational skills, vocabulary/and functions, literature and information, and language writing (NWEA MAP Assessments, 2016). The RIT reading scores, percentile rankings, and Lexile bands were collected and used for both the pre-test and post-test in the study to show reading achievement and growth or lack thereof.

The 2015 RIT values normative study data (NWEA 2015 Normative Data, 2016) provides nationally normed data for both individuals and schools in the areas of reading, language usage, mathematics, and general science. Since the RIT scores are nationally normed, the NWEA MAP assessments also assign a national percentile ranking to each student based on

typical RIT scores for each grade level. For the purpose of this study, NWEA 2015 normative data in the area of reading will be utilized.

> The NWEA 2015 normative data's results are based on K-11 grade level samples. Each sample is comprised of 72,000 to 153,000 student test records from approximately 1000 schools. These samples were drawn randomly from test record pools of up to 10.2 million students attending more than 23,500 public schools spread across 6,000 districts in 49 states. (NWEA 2015 Normative Data, 2016, p. 2).

See the 2015 Normative Data Chart for reading in Figure 4. Also, note the average RIT scores for 2nd grade students throughout the year in Figure 4.

2015 READING Student Status Norms						
	Begin-Year		Mid-Year		End-Year	
Grade	Mean	SD	Mean	SD	Mean	SD
K	137.5	16.78	149.9	13.20	157.6	13.27
1	160.7	13.08	171.5	13.54	177.5	14.54
2	174.7	15.52	184.2	14.98	188.7	15.21
3	188.3	15.85	195.6	15.14	198.6	15.10
4	198.2	15.53	203.6	14.96	205.9	14.92
5	205.7	15.13	209.8	14.65	211.8	14.72
6	211.0	14.94	214.2	14.53	215.8	14.66
7	214.4	15.31	216.9	14.98	218.2	15.14
8	217.2	15.72	219.1	15.37	220.1	15.73
9	220.2	15.68	221.3	15.54	221.9	16.21
10	220.4	16.85	221.0	16.70	221.2	17.48
11	222.6	16.75	222.7	16.53	222.3	17.68

Figure 4. NWEA national normed RIT scores for each grade level.

The qualitative data for the study was collected from teacher and school leader interviews. The interview protocol for teachers can be found in Appendix C, and the interview protocol for school leaders can be found in Appendix D. In order to make sure that the interpretations of the interviews were as valid and reliable as possible, interrater reliability was conducted by giving a sample of the data to my department chair in order for her to review. My department chair also reviewed the coding that took place from the interviews throughout the study.

In addition, an initial observation of both teachers was conducted in order to better inform the research and gain an understanding of the context of small group literacy instruction in the classrooms. These two brief observations were not used for collecting and analyzing research data or to observe students. On the contrary, they were used to guide the development of probing questions to use during interviews and to interpret interview results by focusing on observing the teachers during small group guided reading instruction.

Data Collection

The NWEA MAP Reading assessment data was collected twice during the study, once at the beginning during the winter administration, and lastly at the end during the spring administration. Additional data was collected through administration of the three interviews of each teacher involved in the study, two interviews of each school leader in the study, and the tracking of teacher decision-making within small group guided reading. The data collected was coded in order to protect the identity of the students and teachers involved in the study.

There was a specific order in which the data was collected. Please refer to Figure 7 for a visual representation of this process. First I began by gathering the pre-test data from the NWEA MAP reading assessment. The RIT scores, percentile rankings, and Lexile bands of the students were collected. Then, I conducted an initial observation of both teachers in order to gain an understanding of literacy instruction within their classrooms. These initial observations were focused on the teacher only and concerned with the structure and context of literacy instruction during small group guided reading and the decisions that teachers made within this instructional block. An example of the observation rubric can be found in Appendix B.

Throughout the study teachers tracked their decision making regarding small group guided reading. First, they used their current methods in order to discern the instructional reading levels of the students; such as, their Reading A to Z running record/benchmark tool, NWEA RIT score, and Lexile range. The teachers then noted whether or not the found instructional reading level of each student was within, below, or above the grade level range for 2nd grade students. Refer to Figure 5 for the typical instructional reading level range for 2nd grade used throughout this study.

Reading A-Z	Grade	Lexile Range	Fountas & Pinnell
K	2nd	510L-620L	J
L	2nd	510L-620L	K
M	2nd	530L-810L	L
N	2nd	530L-810L	M
O	2nd	600L-850L	M
P	2nd	600L-850L	M

Figure 5. Reading A-Z correlation chart used for grade level ranges for IRLs and texts. Please note that the 2nd grade within grade level range used through the Reading A-Z program was K-P, which correlated to a Fountas and Pinnell range of J-M. The full correlation chart can be found in Appendix A.

Next, the teachers used their professional judgment to form small guided reading groups, keeping in mind all of the information gathered as well as how many groups could be successfully managed. The students were placed in groups with other classmates whose instructional reading levels were the same, or very similar. Then, the teacher chose the level of text that she considered the best fit for each group and used that text for instruction. For example, the text chosen may have matched some students' found instructional reading levels, but for other students the text may have been either higher or lower than their found levels. Please see Figure 6 for an example of this type of teacher decision making.

The collection of decision making regarding the found instructional reading levels of the students, whether they were within the typical grade level range or not, and the levels of text used with the students during their small group guided reading times were updated into a spreadsheet each time there was an adjustment to one of the areas. The format of the spreadsheet used for tracking was similar to Figure 6, with the addition of dates and additional columns to

track changes made throughout the semester.

Guided Reading Group	Student Number	Found Instructional Reading Level	Found Instructional Reading Level is Within, Below, or Above Typical Grade Level Range	Level of Text Chosen for Guided Reading Group	Level of Text Matches, is Lower, or is Higher than Found Instructional Reading Level
Group A	#1	Level H	Below	Level H	Matches
	#2	Level I	Below	Level H	Lower
	#3	Level G	Below	Level H	Higher
Group B	#4	Level J	Within	Level L	Higher
	#5	Level K	Within	Level L	Higher
	#6	Level L	Within	Level L	Matches

Figure 6. An example of possible teacher decision making in creating guided reading groups.

Additionally, teacher interviews were conducted at the beginning, middle, and end of the study, as well as interviews of the school leaders held at the beginning and end of the study. These interviews occurred in one hour segments in a quiet place that was free from distractions. The location of the interviews was mutually convenient. A consent form was signed prior to the interviews taking place. Each interview was recorded, transcribed and coded. The interviewees received a copy of each transcript in order to check them for accuracy.

The final data collected was the post-test data from the NWEA MAP reading assessment administered at the end of the study. This data included the RIT scores, percentile rankings, and Lexile ranges of each student in the study.

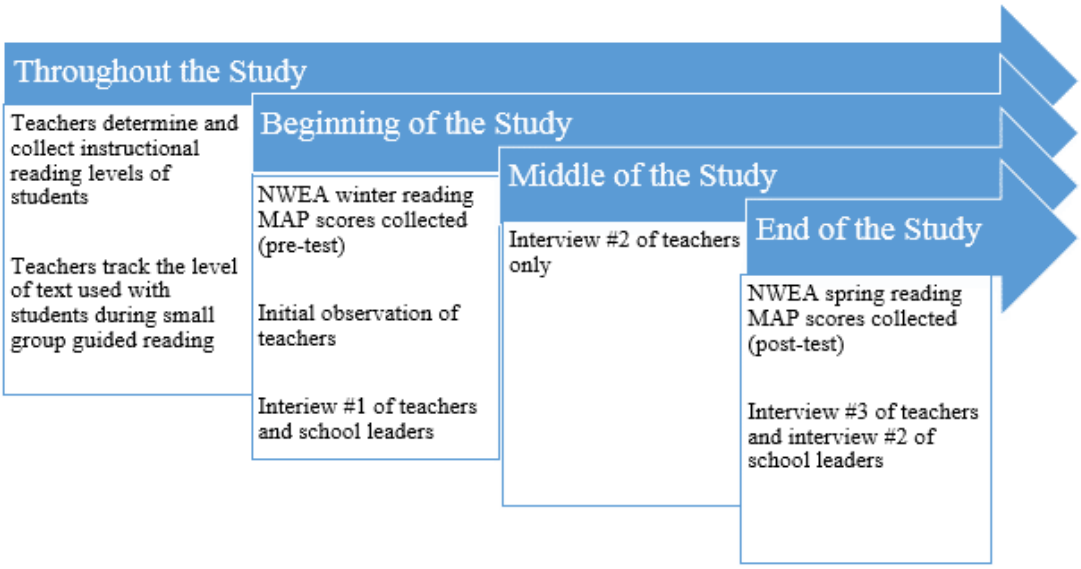

Figure 7. Visual representation of data collection process.

Data Analysis

The data analysis included both quantitative and qualitative methods and tools. Quantitative data was numerical and verifiable, and was able to be statistically manipulated. In order to analyze the quantitative data, both descriptive and inferential statistics were utilized. Qualitative data is the collection of words that describe people's behaviors, actions, and interactions (Roberts, 2010). The qualitative data gathered was in the form of multiple interviews, and both open and axial coding were used to analyze and interpret the data gathered in order to identify common themes and better explain the results of the study.

Quantitative descriptive statistics. I used descriptive statistics to describe and summarize the basic features of the sample. Descriptive statistics helped to describe data through measures of central tendency, such as the mean, to summarize the data collected (Field, 2013).

In addition, I used descriptive statistics in order to compare reading achievement and growth across the student demographics of the students in the study.

Quantitative inferential statistics. Inferential statistics were utilized in order to further analyze the quantitative data, and to assess the strength of the relationships between the independent and dependent variables in the study. For example, in order to determine whether a relationship may exist between the instructional reading levels of the texts used with students and their reading achievement, a Pearson correlation was originally going to be administered between the level of texts used and the student's gain on the NWEA; however, if a smaller number of instructional texts are used, then a one-way ANOVA was planned to be run instead for each level (Field, 2013).

In addition, linear regression with a categorical predictor was planned to be run between the gains on NWEA (dependent variable) and whether or not the texts read matched the found instructional reading level (1), were lower than the instructional level (0), or were higher than the found instructional reading level (2). However, if I do not have a large enough sample size, I had planned on running a one-way ANOVA for all three categorical variables (Field, 2013). The relationship between whether or not a student is reading within, below, or above the typical grade level range and their reading achievement on NWEA was also planned to be analyzed using either linear regression or one-way ANOVA depending on the sample size.

As formerly noted, I had originally planned on using the previously mentioned inferential statistics in my study: Pearson correlation, one-way ANOVA, and linear regression. However, all of these are parametric inferential tests, which make assumptions about the data used that does not fit my data set. Therefore, since my data set was n=34, and most of my data is grouped or

ranked, and not normally distributed, I learned that I needed to use nonparametric inferential tests instead. This knowledge came from my consultation with Dr. James Jones, the Director of Research and Academic Effectiveness at my university. Consequently, I ran the following nonparametric tests for the quantitative analysis portion of my study: Chi-square tests of association, Kruskal-Wallis tests, and the Mann-Whitney Pairwise Comparison tests. These assessments are described in more detail in Chapter 4.

Qualitative open coding. The first step in understanding the qualitative data gathered was to conduct Open Coding, which is also referred to as Initial Coding or First Cycle Coding. The process of Open Coding includes a close reading of the data gathered, followed by breaking the data into distinct sections, and then to compare the data for similarities and differences. In addition, during this initial reading and digesting of the data, initial codes may start to be identified (Corbin & Strauss, 2007). Since my qualitative data was based on multiple interviews, the interview data was first transcribed. Then I examined that data closely through a line-by-line Initial Coding process in order to begin to identify the categories that emerged. The data was then organized into common themes to explain what was occurring (Corbin & Strauss, 2007; Saldana, 2009). Inter-rater reliability was used in order to assure reliable and valid results. This was accomplished through the assistance of my doctoral committee chair who also coded a sample of the interview data.

Qualitative axial coding. Once the Open Coding was completed, then Second Cycle, or Axial coding, was conducted. During Second Cycle coding, the categories and codes were reorganized and synthesized into fewer categories that encompassed broader themes and concepts. Subsequently, I combined the many categories into fewer, more broad categories

during the Axial coding process. During Axial coding, the properties of the categories and subcategories were described and relationships between them were explored (Corbin & Strauss, 2007; Saldana, 2009). Inter-rater reliability was employed during this stage of analysis through the incorporation of my doctoral committee chair who coded some of the information in order to produce a more reliable and valid result.

Limitations of the Study

As in all research, mine had some limitations. There were limitations in the design of the study since an experimental design was not permitted to be completed, and teacher decision making was tracked instead. In addition, the study expanded over the course of a semester based on the district parameters set for NWEA administration, and the approval process necessary to find a school and district in which to conduct the proposed research delayed its launch and prevented it from occurring over the entire school year. Third, the implementation of the study had some limitations based on inherent bias that teachers bring into their instructional practices. In addition, the amount of time students chose to read texts of varying levels outside of guided reading, both at home and at school, was another limitation of the study. Finally, the generalizations of findings from this study may have been affected by the fact that this study was completed within a school district with a relatively limited sample size. Furthermore, another one of the limitations may have been the generalizability due to the study of a single district that was relatively homogeneous.

Summary

This chapter highlighted the research methods that used during this mixed methods study. The research design was QUAN-QUAL in order to incorporate the gathering of student reading data through NWEA, an initial observation of both teachers, followed by three interviews of each teacher, and two interviews of each of the school leaders. In addition, the study considered the level of texts used with students during small group guided reading, and whether or not those texts matched the students' found instructional reading levels, were lower, or were higher their IRLs. The study also took into account whether or not students' found instructional reading levels were within, below, or above the typical grade level range for 2nd grade students. I am hopeful that the results of this study may prove useful in fostering a continued understanding of the complex nature of reading and how we might continue to increase the reading achievement of all students. Chapter Four describes the results of the study.

CHAPTER 4: RESULTS

This chapter describes the results found from collecting and analyzing both the quantitative and qualitative data gathered during this mixed-methods study. I hypothesized that students would both grow and achieve at a higher rate when spending more time reading texts higher than their instructional reading level (IRL) during small group guided reading. Teachers collected quantitative data regarding small group guided reading, including pre- and post- test results, guided reading levels, and time spent reading texts at different instructional reading levels. In addition to one teacher observation per teacher, interviews were conducted with both the teachers and school leaders who participated in the study in order to gain insight into their perceptions. The chapter begins with a review of the purpose and research questions. It is then organized into the following sections: demographics, literacy setting, and participants. Finally, the qualitative analysis of both first and second cycle coding are unpacked and connected to the final two qualitative research questions, and a summary concludes the chapter.

Purpose of Study

One purpose of this study was to determine which instructional reading level text (one that matches, is below, or is above a student's found instructional reading level) used in small group guided reading fostered greater reading achievement and growth as measured by NWEA, and to identify the relationship between whether a student's found instructional leading level was within, above, or below the typical grade level range and their reading achievement scores and growth. The independent variables in this study were gender, race, SES, Limited English Proficiency, special education, the found instructional reading level of students, whether the students' found instructional reading level was within, above, or below the typical grade level

range, and the instructional reading level of texts used with students. The dependent variables were the reading achievement of the students as measured by NWEA RIT and percentile ranking on the pre and post-tests and the achievement growth between the pre- and post-tests.

Another purpose of this study was to determine the perceptions of the teachers as they implemented different levels of texts with the students during small group guided reading instruction. Finally, the perspectives of the principal and professional liaison were sought regarding how they might best support literacy initiatives and how the results of this study might be expanded to other settings.

Research Questions

The following research questions were addressed in this study:

1. What is the relationship between the instructional reading levels of the texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading growth as measured by NWEA?
2. What is the relationship between the instructional reading level of texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading achievement as measured by NWEA?
3. What is the relationship between whether students' found instructional reading levels are within, below, or above the typical grade level range and their end of semester reading achievement and growth on NWEA?

4. What are the perceptions of the teachers as they increase the instructional reading level of texts used with some students?
5. What are the perceptions of the principal and the professional development liaison in expanding the results of this study to other teachers?

Demographics

As a basic overview from the detailed demographic information shared in Chapter Three, this study was conducted in a Kindergarten – 6th grade elementary school with a student population of approximately 410 students in rural school district in the Midwestern part of the United States. According to the Department of Education's website (2016), the elementary school was made up of about 97% White students, with approximately 40% of the population on the free/reduced lunch program. Please refer to Figure 2 for school and state demographic data.

Literacy setting and participants. The study occurred within two second grade classrooms in the same elementary school. The setting of the study was the section of the literacy block dedicated to teacher- led, small group guided reading. Each classroom was comprised of about 17 students each, in which a total of 34 students participated. Both of the teachers, referred to as Teacher A and Teacher B, in this study are in their second year of teaching second grade in this elementary school. Teacher A has had other years of experience teaching kindergarten and the Teacher B is only in her second year altogether. Teacher B had the high ability cluster of students in her classroom. Teacher A had a wide range of students, including two students with IEPs in her classroom. The teachers both willingly collaborated with the professional development liaison last year, focusing on small group guided reading as their primary topic of study. They additionally both excitedly agreed to participate in this study.

The sampling procedure used was a convenience sample based on the professional judgment of the teachers, which was necessary in order for the school leadership to agree to participate in the study. Therefore, this study did not direct teachers to use particular levels of texts with students or to group them in specific ways; it provided a format for tracking those teacher-made decisions in order to analyze the type of impact those decisions may have had on reading achievement and growth. Please refer to Figure 6 for an example of decision making. Finally, the length of the study spanned over an entire semester, which is about an 18-week period of time, ranging from January until approximately the third week of May.

The school leaders in this study were comprised of a building principal and a professional development liaison provided by the local university. The building principal was completing his 7th year at the school during this study. Prior to his role as principal, he taught 5th and 6th grade mathematics at the same elementary school. He did not have an assistant principal, which is one of the reasons that the school partnered with the university. As previously mentioned, the professional development liaison was a literacy professor at the local university who, as part of her position, also supported this school's profession development growth. During the study, the liaison was completing her 4th year in this position with the elementary school.

Findings

The results found through this mixed methods study are organized below by both quantitative and qualitative results. First, the descriptive statistics were created from the quantitative data collected through the study. Next, the inferential statistical findings are described and listed by research question. Finally, the qualitative data from the teacher and school leader interviews were coded using both first and second cycle coding methods. The

dominant codes and major themes are displayed in matrixes and defined in detailed. The combination of unpacking both quantitative and qualitative data provides a more robust understanding of the results of this study.

Quantitative Data

The following section unpacks the quantitative data that was collected and analyzed during this study. Descriptive and inferential statistics were both utilized in order better understand the quantitative data collected and determine if any of the data might be statistically significant.

Descriptive statistics. Teachers in the study were asked to collect data regarding many aspects of student demographics, such as gender, race, Limited English Proficiency, SES, and special education. In addition, teachers collected data specific to each student regarding initial and final instructional reading levels and the amount of time students spent reading a text in small group guided reading that was below, matched, or was above their personal instructional reading level. NWEA Reading pre and post-test results of both achievement and growth were also collected as part of this study by the participants. Descriptive statistics were run and tables were created in order to determine possible relationships among the dependent and independent variables. The mean was used most often as the measure of central tendency to describe the data collected.

Table 1

Mean NWEA Reading RIT Growth and Time Spent Reading Text Higher than IRL of Students.

Amount of Time Reading Text Above (%)	Number of Students	Percentage of Students (%)	Mean RIT Growth	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0-33	9	26.5	5.0	88.9	11.1	0.0
40-67	12	35.3	6.7	100.0	0.0	0.0
73-100	13	38.2	7.2	46.2	30.7	23.1

Note. This table shows the percentage of time students spent reading text higher than their instructional reading level compared to whether their instructional reading level found at the beginning of this study was above, within, or below typical 2nd reading level range in relation to their mean NWEA Reading RIT growth from winter to spring.

As shown in Table 1, the students in Group 3, who read text above their instructional reading level 73-100% of the time, appear to have achieved the most growth as measured by NWEA Reading RIT scores from winter to spring, with an average increase of 7.2 points. The group of students in Group 1, who read text the least amount of time higher than their IRL, displayed the lowest mean growth of 5 points from winter to spring.

In addition, Table 1 indicates what percentage of students entered the study reading above, within, or below grade level expectation. The students in Group 3, who read text above their IRL 73-100% of the time, not only had the highest mean growth, but also had the fewest amount of students who entered the study reading above. In fact, this group of students had a more even spread of percentage of students reading above, within, and below grade level. The fact that the other two groups were comprised of a higher percentage of students who were

reading above grade level, and yet displayed lower mean growth scores is also something to note. However, I did not know if these numerical differences in the means were statistically significant until an inferential statistical analysis was completed.

Table 2

Mean NWEA Reading RIT Growth and Time Spent Reading Text that Matched the IRL of Students.

Amount of Time Reading Text that Matches (%)	Number of Students	Percentage of Students (%)	Mean RIT Growth	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0-13	12	35.5	7.4	58.5	16.7	25.0
20-33	12	35.3	3.9	83.3	16.7	0.0
47-67	10	29.4	8.3	90.0	10.0	0.0

Note. This table shows the percentage of time students spent reading text that matches their instructional reading level at the beginning of this study in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range in relation to their mean NWEA Reading RIT growth from winter to spring.

As shown in Table 2, the students in Group 3 who read text that matched their instructional reading level 47-67% of the time achieved an average growth of 8.3 RIT points, which is more growth than the students in Group 2 who only read text that matched their instructional reading level 20-33% of the time. However, Groups 2 and 3 were both mostly made up of students who were reading above the typical 2nd grade reading level. In addition, the students in Group 1, who only read text that matched their instructional level 0-13% of the time, had a mean growth of 7.4, which is close to the growth made by Group 3 who read texts that matched their IRL 47-67% of the time. The students in Group 3 were more evenly represented

by students who were measured as having instructional reading levels above, within, and below the typical 2nd grade reading level range. Of course until the inferential statistics were completed, as outlined in a section below, I did not know if the differences found in the means would prove to be statistically significant.

Table 3

Mean NWEA Reading RIT Growth and Time Spent Reading Text that was Below the IRL of Students.

Amount of Time Reading Text Below (%)	Number of Students	Percentage of Students (%)	Mean RIT Growth	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0	17	50.0	6.7	76.5	5.9	17.6
7-27	8	23.5	9.3	50.0	50.0	0.0
33-67	9	26.5	3.6	100.0	0.0	0.0

Note. This table shows the percentage of time students spent reading text below their instructional reading level at the beginning of this study in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range in relation to their mean NWEA Reading RIT growth from winter to spring.

The students in Group 3, who read text below their instructional reading level 33-67% of the time, achieved less growth as measured in the NWEA Reading RIT scores from winter to spring. In addition, it is important to note that these students all entered the study reading above 2nd grade reading level. The students in Group 2, who only read text below their reading level 7-27% of the time, displayed the highest mean growth by far at 9.3 points. In addition, the students in Group 1 who didn't spend any time reading text below their reading level, almost doubled

their average growth when compared to the students in Group 3 who read text below 33-67% of the time during the study (Table 3). The differences found in the means had not yet been analyzed through inferential statistics; therefore, I did not know if they would prove to be significant.

Table 4

NWEA Spring Reading RIT at or above 188 and Time Spent Reading Text Higher than the IRL of Students.

Amount of Time Reading Text Above (%)	Number of Students	Met Spring RIT of 188 (%)	Mean Spring RIT Score	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0 - 33	9	100.0	203.8	88.9	11.1	0.0
40 - 67	12	100.0	201.3	100.0	0.0	0.0
73-100	13	69.2	197.0	46.2	30.7	23.1

Note. This table shows the percentage of time students spent reading text above their instructional reading level in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range at the beginning of this study in relation to meeting the spring NWEA goal of 188.

Table 4 indicates that the students in Groups 1 and 2 who read text above their IRL from 0-67% of the time, all achieved their Spring RIT goal of 188 and a higher mean score than the students in Group 3 who read text above their IRL 73-100% of the time. It is important to note that the students in Group 2 all entered the study reading above grade level expectation, and the students in Group 1 had close to 90% reading above grade level expectation, with the other approximately 10% of students entering already reading at grade level. However, the students in Group 3, had a more even split of initial reading level, with about 50% reading above grade

level, approximately 30 % reading at grade level, and about 20% reading below grade level. However, I did not know if these numerical differences in means were statistically significant until I completed inferential statistical analysis, described in the section following descriptive statistics.

Table 5

NWEA Spring Reading RIT at or above 188 and Time Spent Reading Text that Matched the IRL of Students.

Amount of Time Reading Text That Matched (%)	Number of Students	Met Spring RIT of 188 (%)	Mean Spring RIT Score	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0-13	12	75.0	200.0	58.0	17.0	25.0
20-33	12	91.7	197.9	83.0	17.0	0.0
47-67	10	100.0	203.6	90.0	10.0	0.0

Note. This table shows the percentage of time students spent reading text that matched their instructional reading level in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range at the beginning of this study in relation to meeting the spring NWEA goal of 188.

Table 5 shows that students in Group 3, who read text that matched their IRL 47-67% of the time, all met their spring RIT goal of 188 and had the highest mean score of 203.6. However, 90% of these students entered the study reading above grade level expectation. The students in Group 1, who read text that matched their IRL 0-33% of the time, showed that 75% of them met their RIT goal of 188; however, they had the 2nd highest mean growth of 200.0. Until the

inferential statistics were completed, I did not know if these results would end up proving to be significant.

Table 6

NWEA Spring Reading RIT at or above 188 and Time Spent Reading Text that was Below the IRL of Students.

Amount of Time Reading Text Below (%)	Number of Students	Met Spring RIT of 188 (%)	Mean Spring RIT Score	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0	17	88.2	200.2	76.5	5.9	17.6
7-27	8	75.0	196.3	50.0	50.0	0.0
33-67	9	100.0	204.0	100.0	0.0	0.0

Note. This table shows the percentage of time students spent reading text below their instructional reading level in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range at the beginning of this study in relation to meeting the spring NWEA goal of 188.

Table 6 displays that students in Group 3, who read text below their IRL 33-67% of the time, all met their spring RIT goal of 188 and had the highest mean score of 204.0. However, all of these students entered the study reading above grade level expectation. The students in Group 1, who did not read any text below their IRL during small group guided reading, showed that 88.2% of them met their RIT goal of 188; however, they had the 2nd highest mean growth of 200.2. Once again, these results may or may not prove to be statistically significant until inferential statistics are completed.

Table 7

NWEA Lexile Band Increase from Winter to Spring and Time Spent Reading Text Higher than the IRL of Students.

Amount of Time Reading Text Above (%)	Number of Students	Increased Lexile Band from Winter to Spring (%)	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0-33	9	77.8	88.9	11.1	0.0
40-67	12	91.7	100.0	0.0	0.0
73-100	13	84.6	46.2	30.7	23.1

Note. This table shows the percentage of time students spent reading text above their instructional reading level in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range at the beginning of this study in relation to whether or not their Lexile band increased from winter to spring.

Table 7 shows that the students in Group 2, who read text above their instructional reading level 40-67% of the time, had the highest percentage of students increase their Lexile band complexity from winter to spring, at 91.7%. The students in Group 2, all entered the study reading above their grade level expectation as well. However, the students in Group 3, who had a more diverse initial level of preparedness and read text higher than their IRL 73-100% of the time, showed the second highest amount of growth in Lexile, with 84.6% of these students increasing their Lexile band. However, I did not know if these initial findings would prove to be statistically significant until the inferential statistics were completed.

Table 8

NWEA Lexile Band Increase from Winter to Spring and Time Spent Reading Text that Matched the IRL of Students.

Amount of Time Reading Text That Matched (%)	Number of Students	Increased Lexile Range from Winter to Spring (%)	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0-13	12	91.7	58.0	17.0	25.0
20-33	12	66.7	83.0	17.0	0.0
47-67	10	100.0	90.0	10.0	0.0

Note. This table shows the percentage of time students spent reading text that matched their instructional reading level in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range at the beginning of this study in relation to whether or not their Lexile band increased from winter to spring.

As indicated in Table 8, students in Group 3, who read texted in small group guided reading that matched their instructional reading level 47-67% of the time, all increased their Lexile range from winter to spring. In addition, 90% of these students entered the study reading above grade level, and 10% entered reading at grade level expectation. The students in Group 2 had the lowest percentage of students who increased their Lexile band from winter to spring, at 66.7%. These students read text that matched their instructional reading level 20-33% of the time, and over 80% of these students entered the student reading above grade level. Group 1 students only read text that matched their instructional reading level 0-13% of the time, but had 91.7% of the students increase their Lexile. This group of students was also more evenly split between students who entered the study reading above, within, and below grade level

expectation. Of course, the percentages described above may or may not prove to be statistically significant.

Table 9

NWEA Lexile Band Increase from Winter to Spring and Time Spent Reading Text Below the IRL of Students.

Amount of Time Reading Text Below (%)	Number of Students	Increased Lexile Range from Winter to Spring (%)	Student Initial Reading Level Above (%)	Student Initial Reading Level Within (%)	Student Initial Reading Level Below (%)
0	17	88.2	76.5	5.9	17.6
7-27	8	87.5	50.0	50.0	0.0
33-67	9	77.8	100.0	0.0	0.0

Note. This table shows the percentage of time students spent reading text below their instructional reading level in conjunction with whether or not their instructional reading level was above, within, or below typical 2nd reading level range at the beginning of this study in relation to whether or not their Lexile band increased from winter to spring.

As displayed in Table 9, students in Group 2 achieved the highest percentage of students who increased their Lexile band, at 87.5%. These students read text below their instructional reading level 7-27% of the time, with half of the students who entered that study reading above grade level expectation, and the other half reading at grade level expectation. The students in Group 1 who didn't read any text below their instructional reading level in the study, were very close to this percentage, with 88.2% of them increasing their Lexile band. Group 1 also contained students who entered the study reading below grade level, and the other two groups did not. However, I did not know if these numerical differences in percentages were statistically

significant until I completed inferential statistical analysis, described in the section following descriptive statistics.

Table 10

Amount of Students' IRL from Winter to Spring.

Number of Levels Increased from Winter to Spring	Number of Students	Percentage of Students (%)	Mean Percentage of Time Spent Reading Text Higher than Instructional Reading Level (%)	Mean Percentage of Time Spent Reading Text That Matches Instructional Reading Level (%)	Mean Percentage of Time Spent Reading Text Below Instructional Reading Level (%)
1 Level	16	47.0	69.4	20.1	10.4
2 Levels	10	29.4	50.6	30.0	19.4
3 Levels	6	17.6	50.0	34.7	15.7
4 Levels	2	6.0	60.0	23.5	16.5

Note. This table shows the increase in student instructional reading level compared to the amount of time they spent reading texts higher, texts that matched, or texts that were below their instructional reading levels during small group guided reading as measured by Reading A-to-Z benchmark assessments.

Table 10 may indicate that students who spent approximately 50-60% of the time reading text higher than their instructional reading level, about 23-35 % of the time reading texts that matched their instructional reading level, and between approximately 15-20% of their time reading texts that were below their instructional reading level grew from two to four levels in instructional reading level as measured by the Reading A-Z Benchmark System used by the teachers in this study. I was unable to run inferential statistics on this data since the numbers of

students as categorized by how many instructional reading levels they grew were so uneven; therefore, I do not know if they might be statically significant.

Inferential statistics. In this section I have organized the inferential statistics that were run in conjunction with the quantitative research questions in this study. Only nonparametric tests were used due to my small sample size of the data set, at $n = 34$. For each categorical relationship in question, I first ran a Chi-Square test of association in order to determine if there might be any statistically significant relationships between the variables.

When a type of Chi-Square test of association was found to be significant, a Kruskal-Wallis nonparametric test was then run. This test was completed in order to determine if there might be any statistically significant differences among the three groups of students, by comparing the mean ranks, or rank values. This transformation of the data into the rank format was important in order to view my data in a slightly different manner, rather than just looking at the actual means of the raw scores. When the Kruskal-Wallis showed a statistically significant difference among the three groups, then a post-hoc Mann-Whitney Pairwise Comparison test was run in order to determine where the statistically significant differences were found between the specific groups of students.

As previously mentioned, the students in the study where divided into three groups based on a range of time they spent reading a particular level of text. Table 11 shows the breakdown of these three categories and the percentages of time groups of students spent reading these types of texts. The students were placed into groups based on natural breaks in the percentage of time spent reading that particular level of text.

Table 11

Percentage of Time Students Read Different Levels of Text During Small Group Guided Reading.

Students	Reading Text Higher than IRL (%)	Reading Text that Matches IRL (%)	Reading Text that is Below IRL (%)
Group 1	0-33	0-13	0
Group 2	40-67	20-33	7-27
Group 3	73-100	47-67	33-67

Note. This table shows the breakdown percentages of time, divided into three groups, in comparison to reading text higher, text that matched, or was below their IRL.

If my data set had been larger, I would have been able to run logistical regression in order to determine differences between the groups of students based on percentage of time spent reading text higher, that matched, or was below their IRL and whether students met their NWEA projected growth and NWEA spring achievement goal of 188. However, in order to analyze possible significant differences between the groups, the NWEA mean growth and the mean NWEA spring achievement scores of the students were compared to the time they spent reading text at the different categories of levels.

Research Question 1. What is the relationship between the instructional reading levels of the texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading growth as measured by NWEA?

There were no statistically significant correlations found between the amount of time different instructional reading levels of texts were used with students during small group guided reading and their end of semester reading growth as measured by NWEA MAP Reading RIT

scores. However, a related result was found to be statistically significant between the percentages of time students spent reading texts in small group guided reading that matched their instructional reading levels and increasing their Lexile band from winter to spring.

Table 12

Percentage of Time Spent Reading Text that Matched Students' IRL and Lexile Band Increase.

Chi-Square Test	Value	df	Asymptotic significance (2-sided)	Exact Sig. (2-sided)
Likelihood Ratio	6.235	2	.044	.103
N of Valid Cases	34			

Note. Significant at the $p<.05$ level.

As shown in Table 12, the Likelihood Ratio Chi-Square was run and indicated a statistically significant association between the percentage of time students read text that matched their IRL and increasing their Lexile band from winter to spring. The Likelihood Ratio test is a type of Chi-Square test that is sometimes used as criteria for goodness of fit, and builds on the likelihood of the data. This result indicates that the increase in Lexile band and time spent reading text that matched students' IRL is not likely due to chance, at the *p* value of $p<.044$.

Table 13

Kruskal-Wallis Test of the Percentage of Students who Increased their Lexile Band from Winter to Spring and the Amount of Time Spent Reading Text that Matched their IRL.

Total N	Test Statistic	df	Asymptotic Sig. (2-sided test)
34	5.272[a]	2	.072

Note. a. The test statistic is adjusted for ties. Significant at the $p<.05$ level.

Since the Likelihood Ratio Chi-Square indicated a statistically significant result, a Kruskal-Wallis Test was then run in order to determine if there might be a statistically significant difference between the three groups of students. The Kruskal-Wallis Test compared the value of the mean ranks of the three groups of students. As shown in Table 13, once the value of the mean ranks were compared, $p<.072$ was found. The result shows that the differences between the three groups approached significance. Therefore, this may be an area to watch in future studies. Since these results approached significance, a post-hoc Mann-Whitney Pairwise Comparison was completed in order to pinpoint if a statistically significant relationship might exit between the three different groups (Table 14).

Table 14

Mann-Whitney Pairwise Comparison of the Percentage of Students who Increased their Lexile Band from Winter to Spring and the Amount of Time Spent Reading Text that Matched their IRL.

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
2-1	4.250	2.495	1.703	.088	.265
2-3	-5.667	2.617	-2.166	.030	.091
1-3	-1.417	2.617	-.541	.588	1.00

Note. Group 1 read text that matched their instructional reading level 0-13% of the time. Group 2 read text that matched their instructional reading level 20-33% of the time. Group 3 read text that matched their instructional reading level 47-67% of the time. Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. a. Significance values have been adjusted by the Bonferroni correction for multiple tests. Significant at the $p<.05$ level.

As shown in Table 14, there was a statistically significant difference found between Groups 2 and 3 at $p<.03$. As a reminder, the students in Group 2 read text that matched their IRL 20-33% of the time, the students in Group 3 read text that matched their IRL 47-67% of the time. As displayed in Table 8, the students in Group 2 had the lowest percentage of students increase

their Lexile band, when compared to students in Groups 1 and 3. The students in Group 3, who read text that matched their IRL more than the other two groups all increased their Lexile band.

When looking at the adjusted significance information from the completion of the Bonferroni, there does not appear to be a statistically significant difference between Groups 2 and 3. It is important to note that the Bonferroni was calculated in order to control Type 1 errors in my research. However, Dr. James Jones, the Director of Research and Academic Effectiveness at my university, shared that since my data set was so small it is most likely prone to Type 2 errors, and not Type 1 errors. Therefore, he recommended putting more importance on the original levels of significance, and not the adjusted.

Summary of research question 1. When running multiple inferential tests in order to determine if there was a statistically significant relationship between the percentage of time students spent reading texts that were below, that matched, or were above their instructional reading levels and their growth on NWEA MAP reading RIT score, a direct relationship was not found. However, when analyzing the Lexile band growth within NWEA and the percentage of time particular levels of instructional text were used with students, a statistically significant finding did appear (Table 14) between two groups of students when reading texts that matched their instructional reading levels. As shown in Table 8, students in Group 3, who read texts that matched their IRL 47-67% of the time all increased their Lexile band from winter to spring. However, the students in group 2 read text that matched their IRL in small group guided reading 20-33% of the time, and a smaller percentage of them, 66.7%, experienced Lexile band growth during the same time period (Table 8). This finding may suggest that in order for students to

maximize their reading growth, they may want to read text that matches their IRL approximately 50% of the time during small group guided reading. This is discussed further in Chapter 5.

Research Question 2. What is the relationship between the instructional reading level of texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading achievement as measured by NWEA?

Table 15

Percentage of Time Spent Reading Text Higher than Students' IRL and Exceeding/Meeting or Not Meeting the Spring NWEA Reading RIT Achievement Score.

Chi-Square Test	Value	df	Asymptotic significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Linear-by-Linear Association	5.410[b]	1	.020	.018	.015	.015
N of Valid Cases	34					

Note. b. The standardized statistic is -2.326. Significant at the $p<.05$ level.

As shown in Table 15, a Chi-Square Linear-by-Linear Association Test was completed. A linear relationship or trend was shown between the amounts of time students read texts above their RIT achievement score. This finding may be represented by, (χ^2 (1) = 5.410, $p \leq .05$). This result means that the linear association between the percentage of time students read texts higher than their instructional reading level in small group guided reading and their mean spring NWEA reading RIT achievement score is statically significant and was not likely due to chance.

Table 16

Kruskal-Wallis Test of Percentage of Time Spent Reading Text Higher than Students' IRL and Exceeding/Meeting or Not Meeting the Spring NWEA Reading RIT Achievement Score.

Total N	Test Statistic	df	Asymptotic Sig. (2-sided test)
34	7.108[a]	2	.029

Note. a. The test statistic is adjusted for ties. Significant at the $p<.05$ level.

Since the Chi Square Linear-by-Linear Association test indicated a statistically significant result ($p = .05$), a Kruskal-Wallis Test was then run in order to determine if there might be a statistically significant difference between the three groups of students. The Kruskal-Wallis Test compared the value of the mean ranks of the three groups of students. As shown in Table 16, once the value of the mean ranks were compared, a statistical significance was shown to have occurred between the three groups. Significance was noted with a p value of $p<.029$. Therefore, a post-hoc Mann-Whitney Pairwise Comparison test was run in order to determine where the statistically differences were between the three groups of students. This results from this additional test are found in Table 17.

Table 17

Mann-Whitney Pairwise Comparison of Exceeding/Meeting or Not Meeting the NWEA Spring RIT Achievement Score and the Amount of Time Spent Reading Text Higher than their IRL.

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
3-1	5.231	2.411	2.170	.030	.090
3-2	5.231	2.226	2.350	.019	.056
1-2	.000	2.452	.000	1.00	1.00

Note. Group 1 read text higher than their instructional reading level 0-33% of the time. Group 2 read text higher than their instructional reading level 40-67% of the time. Group 3 read text highe than their instructional reading level 73-100% of the time. Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same. a. Significance values have been adjusted by the Bonferroni correction for multiple tests. Significant at the $p<.05$ level.

As shown in Table 17, the students in group 1 read text higher than their instructional reading level 0-33% of the time during the study. The students in group 2 read text higher than their instructional reading level 40-67% of the time, and the students in group 3 read text higher than their instructional reading level 73-100% of the time. In addition, there was a statistically significant difference between groups 3 and 2, as well as between groups 3 and 1. As displayed in Table 4, all of the students in group 2 and group 1 met or exceeded the spring RIT goal of 188, as compared to only about 70% of the students in group 3.

As noted in Table 17, the Bonferroni correction was calculated and indicates that the differences between both of the two groups may not be statistically significant, even though difference between groups 3 and 2 approached significance with this adjusted measure, at $p<.056$. The students in group 2 read text higher than their IRL 40-67% of the time, compared to students in group 3 who read text above 73-100%. However, according to Dr. James Jones, the Director of Research and Academic Effectiveness at my university, since my data set was so

small it is more prone to Type 2 errors, and he recommend putting more importance on the original levels of significance, and not the adjusted. He shared that the Bonferroni is used to control for Type 1 errors that are more likely to occur in large sets of data.

Summary of research question 2. A statistically significant result was found between the instructional reading levels of text read with students during small group guided reading and their end of semester reading achievement on NWEA. To be specific, the significant finding stemmed from the amount of time that students read texts during small group guided reading that were higher than their instructional reading levels and their end of semester achievement of a RIT score of 188 or higher on NWEA at the end of the year. These results may indicate that students might achieve greater results in reading when they read texts above their IRL for up to 67% of the time during small group guided reading, and is discussed further in Chapter 5.

There were several areas in which I conducted statistical analysis that proved to not have statistically significant results. These included

- A Chi-Square test of association to identify whether there was a statistically significant relationship between whether students met/exceeded or did not meet the NWEA Reading Spring RIT achievement score of 188 and the percentage of time they spent reading text in small group guided reading that either matched their instructional reading levels;
- A Chi-Square test of association to identify whether there was a statistically significant relationship between whether students met/exceeded or did not meet NWEA Reading Spring RIT achievement score of 188 and the percentage of time they spent reading text in small group guided reading that was below their instructional reading levels;

- A Chi-Square test of association to identify whether there was a statistically significant relationship between whether students met/exceeded or did not meet their NWEA Reading projected growth from winter to spring and the percentage of time they spent reading text in small group guided reading that was above their instructional reading levels;
- A Chi-Square test of association to identify whether there was a statistically significant relationship between whether students met/exceeded or did not meet their NWEA Reading projected growth from winter to spring and the percentage of time they spent reading text in small group guided reading that matched their instructional reading levels;
- A Chi-Square test of association to identify whether there was a statistically significant relationship between whether students met/exceeded or did not meet their NWEA Reading projected growth from winter to spring and the percentage of time they spent reading text in small group guided reading that was below their instructional reading level;
- A Chi-Square test of association to identify whether there was a statistically significant relationship between the percentages of time students spent reading text higher than their instructional reading level, and whether or not they increased their reading Lexile band from winter to spring;
- A Chi-Square test of association to identify whether there was a statistically significant relationship between the percentages of time students spent reading text lower than their instructional reading level, and whether or not they increased their reading Lexile band from winter to spring;

- A Chi-Square test of association to identify whether there was a statistically significant relationship between gender and both the achievement of a RIT score of 188 or higher in NWEA spring reading assessment and whether growth was met/exceeded or not met on the NWEA reading assessment from winter to spring;
- SES of students regarding lunch status in relation to both the achievement of a RIT score of 188 or higher in NWEA spring reading assessment and whether growth was met/exceeded or not met on the NWEA reading assessments from winter to spring.

In addition, I originally planned on running statistics on Limited English Proficiency, race, and special education; however, there were either no students or just one or two students who would fit into one of these other categories. The majority of the students who participated in this study were almost all Caucasian, spoke English as their first language, and had not qualified for special education; therefore, there were not enough students in these categories to complete these additional analyses.

It is important to note that I had originally planned on using the following inferential statistics in my study: Pearson correlation, one-way ANOVA, and linear regression. However, all of these are parametric inferential tests, which make assumptions about the data used that does not fit my data set. Therefore, since my data set was so small, n=34, and most of my data is grouped or ranked, and not normally distributed, I learned that I needed to use nonparametric inferential tests instead. This knowledge came from my consultation with Dr. James Jones, the Director of Research and Academic Effectiveness at my university. Consequently, I ran the following nonparametric tests: Chi-square tests of association, Kruskal-Wallis tests, and the Mann-Whitney Pairwise Comparison tests to analyze my quantitative data.

Research Question 3. What is the relationship between whether the students' found instructional reading levels are within, below, or above the typical grade level range and their reading achievement and growth on NWEA?

Due to my small sample size and the unequal number of students in the three categories of whether they entered the study reading above grade level, within grade level, or below grade level, I was unable to run any inferential statistics for this research question. However, descriptive statistics were run in order to try to explain any possible connections to the possible impact that their level of preparedness may have had on their achievement and growth. I analyzed descriptive Tables 8 and 4 that are representative of the significantly statistic results found for both research questions 1 and 2. When looking at these tables, it appears as if the students in the groups who all grew in Lexile band, and who all met or achieved the spring RIT goal of 188 predominately entered the study reading above the typical 2nd grade reading level range.

The final two research questions 4 and 5 were answered below the qualitative portion of this section. These questions were designed to be answered with qualitative measures.

Qualitative Data

The qualitative data analyzed below represents the six interviews completed with the two teachers in the study, in addition to the four interviews completed with the two school leaders. The mean interview duration was 38.3 minutes, with the longest interview lasting 67.2 minutes, and the shortest interview lasting 15.7 minutes. The interviews were recorded, scripted in an

electronic format, and then imported into NVIVO software. I also kept anecdotal notes during the interviews in order to determine follow-up questions and start to determine possible themes.

First and second cycle coding. Open coding, or first cycle coding, was completed using NVIVO software in order to determine commonalities and major codes among the responses of all participants (Corbin & Strauss, 2007). Open coding is the process of naming or labeling data in order to help to organize the data into common categories (Corbin & Strauss, 2007; Saldana, 2009). I used multiple methods when conducting first cycle coding. Descriptive coding was used to create nouns that described the data. In addition, In Vivo coding was completed in order to extract words or short phrases from the interviews that the participants used to represent themes, and structural coding was utilized to separate the codes into groups (Saldana, 2009). My dissertation chair assisted me by reviewing my coding of the interviews to minimize potential bias.

Second cycle coding was then conducted in order to refine and reorganize the data codes from the first cycle into conceptual categories and themes. For the second cycle coding, I used axial coding to refine and further develop the codes that emerged from the first cycle coding process (Corbin & Strauss, 2007). Axial coding is the process of relating codes to each other in order to look for basic relationships among the data (Corbin & Strauss, 2007; Saldana, 2009). Once again, I collaborated with my dissertation chair to review the coding that was completed throughout both first and second cycle coding.

Figure 8 displays the total number of interview questions utilized in this study. It further indicates the questions that were asked of only teachers, only school leaders, and which questions were asked of both teachers and school leaders. In order to honor the fact that not all

interview questions were asked of each interviewee type, first cycle coding was applied three times in order to create three different matrixes, which are displayed in Figures 9, 10, and 11. The matrixes displayed in these figures show the common codes that emerged based on the interview questions asked to each type of interviewee. I defined the major codes for each matrix, and included exemplars from the interviews that I felt best represented the code that emerged as dominant. My dissertation chair reviewed my coding of transcripts to increase the coding consistency.

Code	Interview Question Type	School Leader Role	Teacher Role
1: Behaviors Noticed	Teachers only	0	92
2: Behaviors When Instructional Level is Above Grade Level	Teachers only	0	23
3: Behaviors When Instructional Level is Below Grade Level	Teachers only	0	31
4: Behaviors When Instructional Level is On-Grade Level	Teachers only	0	14
5: Behaviors with Easier Text	Teachers only	0	9
6: Behaviors with Harder Text Noticed	Teachers only	0	48
7: Determining Texts Used with Students	Teachers only	7	92
8: Content or Theme	Teachers only	0	19
9: Instructional Level of Text Considered	Teachers only	1	28
10: Chooses Higher Level for Group	Teachers only	0	21
11: Reading Skill	Teachers only	0	35
12: Student Choice	Teachers only	0	5
13: Student Interest or Engagement Considered	Teachers only	0	21
14: Feelings or Thoughts of Teachers	Teachers only	0	51
15: Benefits Encountered or Noticed	Teachers only	1	53
16: Collaboration with Peers	Both Teachers & School Leaders	21	11
17: Considerations for Future Guided Reading	Teachers only	0	16
18: Ideas for Maintaining Focus	Teachers only	0	5
19: Difficulties Encountered or Noticed	Teachers only	6	49
20: Ideas for Future Implementation of Study	Both Teachers & School Leaders	1	2
21: Interactions with School Leaders	Both Teachers & School Leaders	52	9
22: Roles of Instructional Leader	School Leaders Only	35	0
23: School Leader Teaching or Education Experience	School Leaders Only	9	0
24: Structures used by School Leaders to Foster Teacher Collaboration in Literacy	School Leaders Only	32	0
25: Motivation to Innovate GR Strategies	Both Teachers & School Leaders	32	70
26: Noticings for Differentiation	Teachers only	10	158
27: Placing Students in Groups	Teachers only	7	79
28: Flexible Grouping	Teachers only	2	6
29: Frequency of Meeting with Small Groups	Teachers only	3	20
30: Instructional Materials for Guided Reading	Both Teachers & School Leaders	4	44
31: Instructional Methods for Teaching Small Group Guided Reading	Teachers only	7	68
32: Length of time on Texts	Teachers only	2	15
33: Scaffolds Provided to Access Harder Text	Teachers only	0	53
34: Utilizes instructional Reading Level	Teachers only	8	52
35: Tools Used to Determine Instructional Reading Levels	Both Teachers & School Leaders	5	34
36: Teacher Interpretation of Instructional Reading Level Tools	Teachers only	6	33

Figure 8. Interview questions by role.

In addition, Figure 8 displays the coding results from the interviews conducted from both teachers and school leaders, and highlights which questions were asked of each type of interviewee. Three additional matrixes were run in order to better identify and analyze the major codes that emerged from interview question type.

Code	Teacher Role
1 : Behaviors Noticed	92
2 : Behaviors When Instructional Level is Above Grade Level	23
3 : Behaviors When Instructional Level is Below Grade Level	31
4 : Behaviors When Instructional Level is On-Grade Level	14
5 : Behaviors with Easier Text	9
6 : Behaviors with Harder Text Noticed	48
7 : Determinating Text Used with Students	92
8 : Content or Theme	19
9 : Instructional Level of Text Considered	28
10 : Chooses Higher Level for Group	21
11 : Reading Skill	35
12 : Student Choice	5
13 : Student Interest or Engagement Considered	21
14 : Feelings orThoughts of Teachers	53
15 : Benefits Encountered or Noticed	53
16 : Considerations for Future Guided Reading	16
17 : Ideas for Maintaining Focus	5
18 : Difficulties Encountered or Noticed	49
19 : Noticings for Differentiation	160
20 : Placing Students in Groups	81
21 : Flexible Grouping	6
22 : Frequency of Meeting with Small Groups	20
23 : Instructional Methods for Teaching Small Group Guided Readi	68
24 : Length of time on Texts	15
25 : Scaffolds Provided to Access Harder Text	53
26 : Utilizes instructional Reading Level	52
27 : Teacher Interpretation of Instructional Readling Level Tools	34

Figure 9. Matrix of codes from teacher role interviews.

Predominant codes for teachers. As shown in Figure 9, the top three codes the emerged from the responses of the teachers in the study were: *noticings for differentiation*, *behaviors noticed*, and *determining text used with students*. These three major codes have been defined below. The definitions include the criteria used for coding, and are followed by exemplars from the teacher responses in the study. The definitions provided were written specifically for explaining the coding used throughout the qualitative process.

Noticings for differentiation. Differentiation is a teaching technique that is used to individualize instruction for students. Differentiation is when a teacher provides students in the same classroom and/or group different ways to learn, process, and display what they have learned. The content itself may be modified through differentiation. Furthermore, differentiation is also when teachers provide different levels of support to students when processing content, which requires teachers to closely observe and evaluate what students are ready to learn next (need a citation here). The code name of *noticings for differentiation* was used in order to code a response that indicated that differentiation was noticed by the researcher. When I coded the interviews with this theme, I looked for actions and thinking that displayed teachers differentiating their instruction with students. For example, I would code with this theme when I noticed teachers sharing examples of or describing differentiation in their decision making.

When working with one of her students who is reading below 2nd grade reading level, here's what one teacher in the group shared she does to help him access texts that are above his IRL. She stated, "Sometimes I pre-read with him. First thing in the morning I pre-read with him, and then when he comes to the group, he's already had exposure. He's either listened to it or then he reads it with me. "

In addition to pre-reading for differentiation, a teacher in the study added,

> I think I definitely tend to use more of like the kind of graphic organizer-type, getting them to like lay out their thinking and really also put it down into words and not just being able to talk about it, but being able to clearly kind of structure that. Maybe it's into a paragraph for some of my higher ones and maybe just a sentence or two for some. But I think, I think pairing writing with it has kind of been the most beneficial is getting them to put their thoughts like down onto paper. Whether, like that looks differently for each group or each student.

One of the teachers also shared how she differentiated for students who read at different instructional levels. She shared,

> Yes, so that would help him be able to engage in the group with the comprehension because even though his reading level is at a D, his comprehension level should be grade level, because he can do listening comprehension at grade level. So that's how I would kind of support him and a couple others in the group. So they had the same text topic, but it was a differentiated text on butterflies for example, so, so I dropped it down a little bit because it was more challenging for that one group. And then for the groups that I pushed higher at a level, we just did a lot of like, just, you know, like before we were at just a lot of background knowledge discussion, vocabulary and kind of reviewing our text features and kind of, just kind of those strategies just to kind of make sure they had front-load kind of what to expect. They'd also had a lot of exposure to that life cycle and videos and stories and read-alouds and things, so they're very familiar with the content, so that really helped them be successful as well I think.

Behaviors noticed. Behaviors are the actions noticed by the teacher, both verbal and nonverbal, of students while interacting with texts during small group guided reading. This code was used each time a teacher shared something that she noticed regarding student behavior while working with texts during small group guided reading.

One of the teachers shared her observations of when she integrated Readers Theater into her small group guided reading. Readers Theater is when a story is written like a play. The students all have a part to read when they act out the story while reading it, instead of just reading it from a book, they are acting out the story. She stated,

> So we did Dr. Seuss week. We did different Dr. Seuss Readers Theaters, so, you know, *Horton Hears a Who*, which is very complicated text for some students, but they were able to be engaged with that because they could have a part that was maybe a lower reading level. You know, I've got parts for them that were appropriate. So we had, some students doing Green Eggs and Ham, which may have been, I don't remember what level that was…like a J, I think it was listed as a J. But because it was so repetitive and they were so engaged they were able to do that even …(though)it was difficult, even my level D reader could do Green Eggs and Ham, because he practiced so much and we read and we were fluent with that. So, then he also had some small parts in the other one, and so I had more mixed groups those two weeks, because they were doing those two Readers Theaters.

The other teacher in the study stated that she noticed the following behaviors during her guided reading time,

> I think the ones that are reading above their level, I wouldn't call it, I don't really feel like we're seeing frustration; I just think that they might be the ones that, in that competitiveness

> feel like, I'm not where I'm supposed to like…I'm not as far as they are, so I'm going to rush, you know, and then that like pulls back their comprehension. But I mean I've been happy that I haven't seen, I wouldn't at all call it like frustration where it's like to the point of like they don't like to read that hard of a text. I think it's just kind of like that awareness of it's taking me longer. But I do think, I've also seen for my ones that are growing a lot like such a boost in confidence too like even when we sit down and do our like running records and we are talking about, look, we read this level last time and now look where we are, and even when we've gotten this. That's just like one of my favorite things - getting to see that, them being aware of their growth and feeling like then they can go and be a leader in their group, and putting that idea in their head too that it's not just about like you're ahead of them; it's just that like you get to set the example in discussion or set the example in conversations that we have about their reading that we're doing.

Determining text used with students. This code refers to how stories, books, or other texts were chosen to be used with students during small group guided reading. This code was used to designate when a teacher shared something that related to the judgment she used in order to decide which texts to use with students during small group guided reading. The interview scripts were coded with this name when a teacher shared and explained how and why she chose the text, book, or story to use with students during small group guided reading instruction.

Both teachers in the study shared the importance of student choice and interest when determining which texts to use with students in guided reading. One teacher stated,

> I just think I've just always tried to make it like engaging and I think they just like that time with the teacher. I think they're getting attention, and so they feel like that's their time and

> so they're especially enjoying that, and so we've tried to make it fun and like I said that engagement piece … (is important).I've tried to pick things that I know that they will be engaged in, … (for example) when we were doing fairy tales and folk tales for a while, the tall tales. So we did a John Henry Readers Theater, a Paul Bunyan and different level texts. Now our last couple of weeks we've been having non-fiction again … (because) we had a little delay from non-fiction. But, I've tried to choose topics that they were interested in and it makes a difference when they can do some of the text selection. I can have them pick what text we do the next week out of a level basket, and that helps. I do notice, I have one student I feel is one of my students, and probably, we just did our high ability testing and I suspect that he could be gifted in math and if not in reading as well. But, you know, he is one that as the text has gotten harder, more difficult, we've asked more text-dependent questions. He just wants to tell you what he knows, like for him to have to rely on the text has been a little bit frustrating and he kind of acts like…that's the first time I've had any kind of, not flack at all, but just kind of like … (difficulty with a text). With my higher group, we've been working on text organizations and we've been reading short passages and they've had to sort by how it was organized, and, you know, he struggled with that a little bit the first couple of days, and it was out of his comfort zone. So for him to have to think about, is this a cause and effect passage? So he really just wanted to rationalize his thinking. Yes. So, not that he was not engaged, but it was just kind of like, oh, I've given him something that's difficult. That's probably one of the first times, you know.

In addition, teachers both discussed their use of gaining instructional reading levels and using these as a part of their decision making process when determining texts to use with students. The other teacher captured this common thinking well we she stated,

> I think definitely letting the Reading A to Z level at the instructional level that we found has kind of determined those groups and kind of kept them pretty steady. I think since we met right off of there, I think there was like one switch maybe, or maybe that was even right before we met when we actually found the levels, I think a couple kids switched. But for the most part, it stayed pretty steady. They just have all grown together for the most then like we were doing even when we first met, the, kind of picking that, like, almost like the highest one, like the highest level from that group and then finding a text.

Code	School Leader Role
1 : Roles of Instructional Leader	42
2 : School Leader Teaching or Education Experience	9
3 : Structures used by School Leaders to Foster Teacher Collaboration in Literacy	34

Figure 10. Matrix of codes from school leader interviews.

Predominant codes for school leaders. The top two codes that emerged for school leaders (Figure 10) were*: roles of instructional leader* and *structures used by school leaders to foster teacher collaboration in literacy*. These two major codes for school leaders have been defined below. The definitions include the criteria used for coding, and are followed by exemplars from the school leader responses in the study. The definitions provided were written specifically for explaining the coding used throughout the qualitative process.

Role of instructional leader. For this study, the term role of instructional leader refers to the job that school leaders played in the development and growth of teacher instructional practices in the area of literacy instruction, particularly small group guided reading. This code

was used to designate when a school leader shared something about his or her role that was connected to the professional development of teachers.

The principal of the school who participated in this study had the following to say about the role he plays as instructional leader of the school. This quote captures the common theme of his responses regarding instructional leadership.

> Yeah, I think one of the important things for us is giving teachers the power, the authority, the opportunity to explore and try out, you know, different options. This is a great example with this study. So I've got two teachers that are involved in the study, Teacher A and, Teacher B who are outstanding teachers and what I see my role in that process is allowing them to, you know, take a look at that data, dive into some different options, explore those and then because of the type of teacher leaders that they are, then they're able to perpetuate some of those ideas out to the rest of the building and some of their colleagues. So my role in that process is just that the role of a facilitator. So giving them the support they need, giving them the time, the resources they need. (As well as)… keeping it within the confines of what's going to be appropriate for our school. So not complete free rein, but, honestly, just putting some bumpers, if you will, out there and letting them kind of run with the ideas as they see fit. And then once we see something that's positive, being a cheerleader in that process, an advocate, if you will, for what they're finding and some of the things that we now can implement for our school and our students.

The university professional development liaison shared how she is both a facilitator and collaborator when it comes to her role as an instructional leader in the building.

> Most of my work as a professional development liaison has been really driven by what the teachers want and feel that they need as far as professional development. This is my fourth year, so it has taken a few years to just first build that relationship and that trust with the teachers and the entire school. I start out each year setting goals as part of the professional development school structure that's in place with the state, so we set goals, both for supporting the university students, as well as supporting teacher professional development and K-12 student learning are kind of the three main areas. So, when we have those meetings each year, annually, there are usually a few kind of professional development things that teachers are interested in learning more about. I've done work with bringing others in to help support professional development.

Structures used by school leaders to foster teacher collaboration in literacy. For this study, this code was used to determine aspects of the interviews that described the ways in which the school leaders created structures or processes to encourage teachers to share ideas, data, new learning, benefits, and struggles within the subject of literacy instruction. The term collaboration in this code refers to teachers working together in order to share ideas, solutions, and wonderings.

The principal in this study spoke about his role in developing and encouraging teacher collaboration to increase the skills of his teacher in general, but also in the area of literacy instruction.

> You know, this first thing is going to seem very simple, but it wasn't in place prior to me being here, and that is common prep periods. It seems very simple in theory, but it really does make a big impact in some of the sharing that takes place. So every grade level is

> devoting one day a week to just coming together for professional sharing. Now that could be lots of things. It's not so structured that they have, you know, a talking format: Well, what's hot today? What are things that we need to talk about? What are concerns? And it could be a student issue all the way up to a big-time curriculum issue. So I think that's one of the disciplines that allows that to happen. Then, naturally, as you know, teacher leaders step up in that process and take different leads. By being the primary evaluator for everybody in my building, I'm able to see that both from an evaluation standpoint, but just more even as an observation standpoint. And so oftentimes those same teacher/leaders will turn around and present to other teams or present at staff meetings, or cross grade level. Now all of a sudden we're going to bring K-12 together and we're going to sit down and have a discussion about this. And so people like Teacher A and Teacher B will bring ideas to me and say, hey, here's a gap. I think we can solve this with Raz-Kids. What do you think about pushing this out to everybody? Alright, let's talk about it. Great. How are we going to train everybody? Well, I've already trained on it. Let's do it and let's push that out. So I'm blessed to have a lot of teacher/leaders that kind of run ideas by me to say, what do you think, yeah, let's run with that. Or, hey, here's an idea we can't run with. But through that collaborative process, teacher/leaders do step up and I know who those people are in the building and I know who to go to if I need an idea supported, changed, enhanced, whatever. Whatever you may say.

The professional development liaison, had the following to say regarding the type of structures she has utilized with the staff in her efforts to provide professional development and opportunities for the staff to collaborate.

> And then last year, rather than having a goal for everyone, we kind of asked teachers what study groups they might be interested in taking part in. And there was interest in the lower grades, K-2. There were a couple of second-grade teachers who were really interested in learning more about how to structure their guided reading lessons. So we did some informal study groups using, is it Richardson's, Next Steps in Guided Reading. We used kind of that text and we just met informally. We would look at different portions of that text and then coming together about monthly after school to discuss how things were going, talk through issues and really just let it be. I was more of a facilitator and helped facilitate the study group conversations with those teachers.

Code	School Leader Role	Teacher Role
1 : Collaboration with Peers	22	11
2 : Ideas for Future Implementation of Study	2	3
3 : Interactions with School Leaders	53	9
4 : Motivation to Innovate GR Strategies	33	70
5 : Instructional Materials for Guided Reading	4	44
6 : Tools Used to Determine Instructional Reading Levels	5	34

Figure 11. Matrix of codes from both the teacher roles and the school leader roles.

Predominant codes for school leaders and teachers. In analyzing Figure 11, it is important to note that the teachers in the study each completed three interviews, at the beginning, mid-study, and at the end. The school leaders only had two interviews: one at the beginning and the other at the end of the study. Therefore, when analyzing the results of the questions that were asked of both teachers and school leaders, *motivation to innovate GR strategies* was the one major, common code that emerged among both groups. The definition of this code is below and includes the criteria used for coding, as well as exemplars from both school leader and teacher

responses in the study. As previously noted, the definition provided was written specifically for explaining the coding process used in this study.

Motivation to innovate GR strategies. For this study, this code has to do with what motivates teachers and school leaders to try new things within the area of small group guided reading. The initials GR refers to guided reading in the code. In addition, as I read through the scripted interviews I also used this code to delineate when teachers and school leaders shared observations and future thinking that resulted from areas new practices in guided reading.

The teachers in the study shared many common reasons for why they are motivated to innovate, or try new things, in the area of small group guided reading. One of the teachers in the study captured this thinking well when she shared the thoughts below.

> I think I see in my kids so much when they succeed, especially at something that's challenging or new to them, because that's kind of difficult for them. So when they actually do …(persevere) and they're so excited and so proud of themselves, not for doing something that's easy to them, but for doing something that's hard and that they overcame, I think that that pushes me, …(because) I know that, OK, I'm responsible for them learning that. They're not just going to know to challenge themselves. They're not just going to know to keep trying or overcome that, and so I think when I kind of step back and realize like I'm their one second-grade teacher, and so I'm the only one this year that gets to do that for them. And so, just I guess they motivate me to, I want the best for them and so I think about myself and think, OK, well, am I doing the best teaching I can do when I get too comfortable with something, and they get bored and I get bored? That's the time when I'm like, nope, then we have to try something different and we have to read something

> different or, you know, just take a break from that. And so I think just thinking about how I like to teach and how I want to learn is … (that) I want to be engaged with my teaching just like they want to be engaged with their learning. And so just kind of thinking about……(if) I'm asking them to do something that I wouldn't want to do, that wouldn't be productive for me to do when I'm learning...(so) then I just don't …(do that).

The school leaders in the study also had a student focus as part of their motivation to innovate, or to encourage innovation in their teachers. The professional development liaison had a slightly different perspective, since she views the teachers in the school as her students in the quote below:

> I'm sorry… what motivates me to encourage innovation? I think, and I may be off on the kind of not getting exactly what you're asking here, so stop me if I am. But I think being in classrooms, particularly the most time I spend in classrooms is observing my students, my pre-service teachers as well as my student teachers, so being in classrooms and seeing that, seeing that innovation that's going on encourages me to promote more of that, seeing kind of the possibilities collaborating with my colleagues on faculty and literacy and things that they're doing. We do a lot of collaboration, sharing of what we're doing in our classrooms, what we've seen out in the field and the amazing results that they've had with various implementation of, an implementation of technology, different ways of structuring guided reading, for example, we've been talking about recently, yeah.

The school principal shared his thinking on why he encourages innovation in his classroom teachers in all areas, which included literacy instruction and small group guided reading.

> Well, first and foremost I think the challenge of the student every year is that much harder. The amount of baggage that students bring with them on a daily basis is ever increasing, and I'm not speaking out of emotion, I'm speaking out of fact. So with that being said, we've got to find new and creative approaches to reach them and so if we're sticking with the old standard and we're not trying to get any better, we're certainly declining. So everybody, me being the first person, has a role to play in trying to get better on a daily basis. And, so, that's much of me as a cheerleader or a, and a curricular leader to make sure that we are doing that, and providing like we talked about the mechanisms by which that happens. Yeah, if we're not getting better, then, you know, that's a real concern. We're three A's back to back to back, and that's because we're trying to get better on a daily basis.

Second cycle coding. During second cycle coding, and as previously noted, I used axial coding determine the relationships between the codes in order to determine broader conceptual categories or themes (Corbin & Strauss, 2007; Saldana, 2009). The main categories that emerged from second cycle coding were the themes of *change process*, *educational leadership*, and *pedagogy* (Figure 12). These themes emerged from the rereading of the qualitative data and utilizing the axial coding method that focuses on determining how the codes are related. Once the three themes were established, the interviews of all participants were then recoded using these three main themes, which are defined below, along with exemplars from the interviews. Once again, my dissertation chair reviewed my second cycle coding in order to increase coding consistency, and to decrease the possibility of bias.

Conceptual Category	School Leader Role	Teacher Role
1 : Change Process	16	38
2 : Educational Leadership	21	5
3 : Pedagogy	18	71

Figure 12. Three main themes from second cycle coding.

As shown in Figure 12, the most predominant conceptual category that emerged from teacher responses was the theme of *pedagogy*. In contrast, the main conceptual category that emerged from school leader responses was the theme of *educational leadership*. In addition, it is important to note that the teachers in the study each conducted three interviews, at the beginning, mid-study, and at the end. The school leaders only had a beginning and end of the study interview, without a mid-study option. Therefore, the conceptual category called *change process* was the one common theme that materialized among the responses of both school leaders and teachers during the second cycle coding process. The dominant conceptual categories that emerged separately for teachers and school leaders have been defined below, along with the one category that both groups found in common. The definitions below each theme include the criteria used for coding, and are followed by exemplars from both teacher and school leaders in the study. The definitions provided were written specifically for explaining the coding used throughout the qualitative process (Figure 12). The second cycle coding process was overseen by my dissertation chair for coding consistency and the reduction of bias.

Predominant category for teachers. The following is a description of the predominant category that emerged from teacher responses during axial coding.

Pedagogy. For this study, *pedagogy* refers to the method and practice of teaching in education, and how the method and practice of teaching influences the growth and development of learners. This code was applied when teachers or school leaders shared something in their interview that was rooted in pedagogy. By this I mean, anything that shed light on the decision making, the processes, the procedures, or the professional development that impacted the choices they made, along with impact noticed on students.

Both teachers discussed the thinking and decision making that they used to determine the instructional reading levels of their students, as well as what text to use with them in small group guided reading. The teachers both noted that they used a combination of a couple of different tools in order to determine the instructional reading levels of their students. One of the teachers shared her thinking below, which was also echoed in the other teacher's thoughts throughout the interviews.

> We use the Raz-Kids, the Benchmark books, so levels A through Z … (which) comes with their running record part, and then there's also comprehension questions too. So we kind of let that take an effect of moving on to another level or finding their instructional …(reading level).We've let NWEA kind of help with that a little bit, but the Raz-Kids was definitely the most beneficial to sit down and do their Running Records. They also have passages for each level, too, that I think when it comes to doing… (these) throughout the semester, that would be the kind of thing … (use for) progress monitoring. But to find their levels, we just use the books.

The other teacher interviewed provided some additional thinking that highlighted more reasons that represented why they both choose to conduct running records and teach small group

guided reading. In addition to the level of accuracy and comprehension to determine the level of text to use, the teachers also considered student interest and engagement when determining text to use during small group guided reading

> I just couldn't imagine teaching reading any other way. So, mostly, I just think it lets you keep such an accurate pulse on where every child is. Like I just, I feel like that's the only way you're going to know who are they are as readers, where are they struggling, what do they need next? What works for them? What are their interests? What engages them? I mean, there's just, that information just can't come any other way. So even when someone offered to do the running records, you know, I'm like no thank you - if I'm going to do that, I want to do it, because…that's the most valuable piece for me is seeing kids interact with text is the most beneficial part.

Predominant category for school leaders. The main category or theme that was evident amongst the responses of the school leaders in this study has been described below. There is both a definition provided, along with some exemplars that best match the theme.

***Educational leadership*.** This code refers to working with and guiding of teachers toward improving educational practices and refers to positions in a school that have leadership roles. The role of *educational leadership* is typically rooted the positions such as a principal, assistant principal, professional development liaison, department chairs, or academic deans. For this study, this code was applied when a teacher or school leader shared something that referenced the impact, importance, role, or involvement of educational leadership.

A common message among both of the school leaders in the study was that they both felt as if their main roles in instructional leadership and development were first grounded in their

relationships with the teacher. They also both felt that it was important for the teachers to want the professional development in order for it to be productive. Therefore, they both shared that they were more facilitators that paved the way for teacher growth and development through giving them permission to try something new and provide resources. The professional development liaison shared the following regarding her role in the building in the area of educational leadership.

> Yes. So as the PD liaison, I go into the schools and as I told you before, I have a limited amount of time in there. Only about three hours a week is dedicated to my time in there, so what I'm able to do is limited and I have tried to spend a lot of time building relationships with the teachers, and so it's just been the past year or so that I've actually had those relationships in place to be able to try to implement some of the PD. So I have, in collaboration with this school, each year we meet and talk about their goals and how the university might help to, to kind of meet some of those goals. And one of the goals that the school had was to improve reading instruction especially in the primary grades. So last year, I worked with a group of self-selected teachers who were interested in learning more about guided reading. We did an informal study group.

The principal in the study expanded on his philosophy of *educational leadership* and role the he plays as the main instructional leader of his school.

> Yeah. So my role in that is similar to how I described it in supporting the teachers is to bring some of those new ideas and some of those ideas may come from our teachers in different areas. Some of them may come from outside research; some of them may come from, you know, conferences, what have you. But then putting that in the hands of the

> teachers and letting them run with it. So, hey, here's a great idea. Here's a great new research tool. Let's explore this and let's run with it. I don't see my role in that process of dictating: you've …(got to) do this, you've …(got to) do that, you've …(got to) do this. But it's more of an opportunity to give them the tools, open their experiences to the new ideas and then let them run with them. Once we find it to be successful over the test market of certain classrooms, then it becomes more, hey, this is something we've got to do and this is the way we're going to put this into practice in our school. So it's still that role of a facilitator for me.

Predominant category for both teachers and school leaders. The predominant category that materialized for both the teachers and school leaders in the study has been defined below. The description contains a definition used for the purpose of axial coding in this study, as well as exemplars that highlight the theme.

Change process. The overarching category of *change process* in this study refers to the acceptance, adaptation, and implementation of change by people in a school setting (Guskey, 2002). This code was applied when a teacher or school leader shared something they were trying that emerged from some type of professional development, if they were implementing a new idea, or changing a practice. This included observing the impact of the new practice, noticing a difference on student achievement or behavior with the change, and his or her beliefs about the new practice.

The teachers in this were engaged in the close monitoring of the level of texts used with their students in small group guided reading. When sharing things they noticed that were changes to their pedagogy, both teachers ended up sharing that they found that there was benefit

to students spending time reading texts at different levels. This was a surprise to them. They found it to be important to spend time reading text above, that matched, and that was below their reading level. One of the teachers indicated the following:

> I think I've seen the benefit of, it's been interesting to be really focused on like, where are they? Are they reading higher or lower or matching their level? I think that variation of some of them doing that sometimes, and then sometimes we're higher and sometimes we're lower, I think it's been, it's just been interesting to see. It's kind of hard to keep track and I think when you're not really focused on that aspect of it, and you're always just pushing, pushing, pushing, I think that …(tracking) kind of helped me to, …(say) OK, they don't always have to be like pushed, pushed, pushed. I think there's room for whole group and there's room for kind of bringing it back and letting them do something that's on their level or maybe lower or maybe, you know, and kind of …(like) that.

The other teacher in the study echoed similar thinking about using multiple levels of texts with students during small group guided reading.

> Yeah, I think that, I like a variety of text levels. I like that sometimes they're reading above their instructional level. I like that sometimes it was, everyone was reading the same level, …(because)that was really your true measure of, this is second-grade material. How can they interact with it?

When asked about his role as an instructional leader of the school, the principal shared that change in instructional practice is rooted, not in just the professional development itself, but on seeing the impact that the new practice has on students. This was what he shared about how change typically manifests in his building.

> It's definitely the freedom to try something new and it's also the opportunity to kind of share that as well. So we've got avenues, whether those just be classroom observations all the way up to, you know, more of a formal presentation at a professional day, or, you know, a staff meeting: Hey, these are some new things I'm trying. Here are the reasons I'm trying them. I think we ought to really explore these and put these to practice as well, and not be afraid to say, this worked really well and here are the few things that didn't work well. So, when a teacher stands up and shares those things, the buy-in is often greater than if I would stand up as someone a little bit disconnected in my role and try to present that. So, you're exactly right. The freedom to fail. The freedom to try. The understanding that there's no pressure for us to feel like, if this doesn't work there's going to be repercussions or anything like that. We know that the decisions we're making are definitely goal-oriented for the best interest of the students. And if it doesn't work, that's OK too.

In addition, and more directly related to this study, the principal shared a change that they implemented based on this study.

> Well, I think that the two that were involved already are seen as teacher leaders in the building, but being a part of this project, that was kind of able to add to their leadership repertoire, if you will, and from a nuts and bolts standpoint, that does affect their evaluation. So from that, they got credit for above and beyond work that they, that they did. But moving forward, we recognized that some of the guided reading levels and the identification of those was a real concern, and so to the early literacy grant, we've now purchased some software that's going to aid all of our K-2 teachers in gathering that much-

> needed data as we, you know, here we were in year one of NWEA and we said this is a weakness of ours. So I think that's one thing that came out of this study to say we've got to have some additional tools. That NWEA's great, but we need to really hone in on that guided reading level and guided reading instruction and in order to be able to do that, then this software piece is going to, you know, be necessary. So, I don't know that that directly came out of the two of them, but it certainly was an outcome of the study for us and it's kind of a "ah-huh", if you will.

Research Question 4. What are the perceptions of the teachers as they increase the instructional reading level of texts used with some students?

Throughout the interviews teachers shared behaviors that they noticed in students and instructional methods that they utilized when the level of texts were increased during small group guided reading. One of their perceptions was that they needed to differentiate in order for more difficult texts to be read successfully. These perceptions were highlighted within the code of *noticings for differentiation*. In summary, both teachers shared that they utilized graphic organizers to incorporate writing into reading and that they developed background knowledge through vocabulary work, videos, and the reading aloud of other texts on the same topic. They shared that these were some of the structures that they provided for the students who were reading text above their instructional reading level in order for them to be successful.

Furthermore, within the code of *behaviors noticed*, the teachers included additional thinking about what occurred when students were reading text above their reading level. One of the teachers noted that when she integrated Readers Theater into the reading of above level text, the students were able to access it because of the repetition needed to perform it, and the level of

engagement and excitement the students portrayed. The other teacher in the study shared that some of her students wanted to rush through the harder text due to competitiveness, and that could limit their comprehension if not caught. In addition, she shared that she saw a lot of growth in her students who were being pushed and that their confidence as readers increased as they owned their data and celebrated their growth. Finally, teachers both also shared that when their students succeed at something difficult or challenging, such as reading a text higher than their instructional reading level, it motivates them to become better teachers. This is highlighted in the code called *motivation to innovate GR strategies.*

Research Question 5. What are the perceptions of the principal and the professional development liaison in expanding the results of this study to other teachers?

There were no major findings in my qualitative research that directly answered this question. However, within the conceptual category, or theme, of *change process* the principal in the study stated that they were going to implement a change based on their involvement in this study. He reflected that through this study they realized that they needed a better curriculum or program for identifying instructional reading levels and instructing small group guided reading. He decided to use early literacy grant funds to purchase an online reading assessment program called Raz-Plus, which is connected to Reading A-Z, that will provide instructional reading level data and that also provides a library of online leveled texts. While this is not his perception of expanding the results of this study, it is a change that has come from his perception of a change that needed to be made through participation in the study.

The classroom observations completed at the beginning of the study were intended to gain a more holistic perspective of the context and structure of literacy instruction in these

environments and were not intended to answer a research question. However, as I reviewed the rubrics that I completed based on the observations, a couple of themes emerged. First of all, I was able to observe the reading model being implemented in both classrooms. For example, the teachers were grouping students for small group instruction by similar instructional reading level. In addition, the teachers were using texts higher than some of the students' instructional reading levels in the groups that I observed. Finally, I witnessed both teachers utilizing scaffolds with their students through the use of graphic organizers to concretely track the use and results of reading strategies, repeated readings, pre-teaching vocabulary, and listening in to students read individually in order to coach and provide support as needed. Based on the findings from these observations, I am confident that the teachers understood the concepts that I was striving for and that they were implementing this study with some fidelity.

Summary

In this chapter, I presented the results of my quantitative and qualitative research from this mixed methods study. One of the major quantitative findings showed that there was a statistically significant relationship found between the percentages of time students spent reading text higher than their instructional reading level, and whether they met or did not meet the NWEA spring RIT achievement score of 188. This statistically significant result may indicate that students who read text higher than their instructional reading level during small group guided reading up to 67% of the time have achieved greater achievement in reading.

Although not directly related to one of research questions, there was also a statistically significant relationship found between the amount of time students spent reading text that matched their IRL and whether or not they increased their Lexile band in NWEA. This result

may indicate that students who read texts that match their IRL 47-67% of the time may have more growth in reading; however, it also may indicate that students who only read text that matched their IRL 13% of the time or less may also achieve good growth in reading.

From the qualitative data collected and analyzed in the form of interviews of the two teachers and the school leaders in the study, along with my anecdotal notes, some main codes and major themes emerged. During first cycle coding, the main codes that emerged for teachers were *noticings for differentiation*, *behaviors noticed*, and *determining text used with students*. One of the commonalities found among both teachers was that they both put scaffolds put into place for students to be able to successfully access higher texts, and that with these scaffolds they saw confidence in their students grow, along with some competitiveness since the texts required more time to read. Another main similarity was that both teachers used a mix of students' instructional reading levels and their interest when determining which text to use with the groups. In addition, when given the choice between which levels of text to use with each group in small group guided reading, they typically chose the highest instructional reading level represented.

On the contrary, the codes that emerged for the school leaders were *roles of instructional leader* and *structures used by school leaders to foster teacher collaboration in literacy*. A major common finding in both school leaders was that they both felt as if their main purpose as an instructional leader was to empower and support teachers in the areas of professional development that the teachers felt was important to focus on. Throughout the first cycle coding process, there was also one common major code that emerged for both teachers and school leaders; *motivation to innovate GR strategies*. The teachers and school leaders in the study all

shared that helping students be successful in school and life was what motivated them to innovate or to encourage innovation in others.

Finally, during second cycle coding, three main themes emerged, they were *change process*, *educational leadership*, and *pedagogy*. The one common theme amongst teacher and school leaders was *change process*. The teachers in the study both shared that since they have been monitoring the level of text used with students throughout this study, they have both seen a high value in using multiple levels of texts with their students in small group guided reading. They specifically spoke of benefits in using text that was higher than students' IRL, that matched their IRL, and that was below the IRL of students during small group guided reading. The principal shared that change in his building typically takes place when teachers have tried something new and have seen a positive impact on students. In conclusion, the full results of both the quantitative and qualitative findings from this study, along with their possible implications, are discussed further in Chapter 5.

CHAPTER 5: CONCLUSIONS

This research was conducted in order to consider whether students are achieving their maximum growth and achievement in small group guided reading by reading texts that match their found instructional reading levels, which is universally accepted as best practice in reading instruction. This mixed methods study employed both quantitative and qualitative research methods, and was conducted over the course of one semester.

This chapter begins with a summary of the study, and highlights the conclusions drawn from the data collected and analyzed in chapter 4. The major findings are explained in-depth underneath each research question to which they pertain. The findings are then looked at through the lens of the research literature that has been collected on the topics in this study. Finally, the chapter concludes with surprises, implications for future practice, and recommendations for further research.

Overview of the Problem

The placement of students in small guided reading groups at their instructional reading level is universally accepted as best practice by the majority of educators, reading specialists, and interventionists (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996.) For decades, Emmett Albert Betts (1946) was considered the forefather of how to determine a student's independent, instructional, and frustration reading levels through his work on informal reading inventories, or IRIs (Boley & Pennock 1975; Ekwall, 1976; Johns & Magliari, 1989; Pikulski, 1990; Powell & Dunkeld 1971; Williams, 1959). However, there are some scholars questioning the origin of the research data used regarding the percentages recommended for the reading levels set forth in

Betts' 1946 book, *Foundations of Reading Instruction* (Cooper, 1952; Johns & Magliari, 1989; Powell & Dunkeld, 1971). More recently, according to the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), the level of text complexity must be raised for all grade levels in order for the United States to compete internationally. However, would raising this bar of text difficulty close the gaps in reading achievement on its own or do we also need to raise our expectations of what constitutes an instructional reading level? Is the percentage range for which we base an instructional reading level accurate and rigorous enough? Will students become frustrated and achieve at lower levels if the text is too difficult for them? These are the types of questions that propelled this study. Furthermore, there have been minimal studies conducted on the effect of the level of text used with students during small group guided reading and the subsequent impact on reading achievement and growth.

Purpose of the Study

The main purpose of this study was to determine which instructional reading level text (one that matches, is below, or is above a student's found instructional reading level) used in small group guide reading fostered greater reading achievement and growth as measured by NWEA, and to identify the relationship between whether a student's found instructional reading level was within, above, or below the typical grade level range. The independent variables in this study were gender, SES, race, Limited English Proficiency, the found instructional reading level of students, whether students' instructional reading levels were within, above, or below the typical grade level range, and the instructional reading levels of the texts used with students during small group guided reading. The dependent variables were the reading achievement and

growth of the students as measured by NWEA MAP RIT and nationally normed percentile on the pre and post-tests and their Lexile band growth also found within NWEA.

Another purpose of this study was to determine the perceptions of the teachers as they implemented different levels of texts with the students during small group guided reading instruction. Finally, the perspectives of the principal and professional liaison were sought regarding how they might best support literacy initiatives and how the results of this study might be expanded to other settings.

Research Questions

The following research questions were addressed in this study:

1. What is the relationship between the instructional reading levels of the texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading growth as measured by NWEA?
2. What is the relationship between the instructional reading level of texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading achievement as measured by NWEA?
3. What is the relationship between whether students' found instructional reading levels are within, below, or above the typical grade level range and their end of semester reading achievement and growth on NWEA?

4. What are the perceptions of the teachers as they increase the instructional reading level of texts used with some students?
5. What are the perceptions of the principal and the professional development liaison in expanding the results of this study to other teachers?

Review of the Research Methods

This study utilized a mixed methods research design with the hopes of honoring the multifaceted, complex process of reading achievement and growth. As is common, this mixed methods approach study consisted of both quantitative and qualitative measures, and attempted to give more breadth and depth to the research by enabling the questions of both *what* and *why* to be addressed when analyzing and explaining the data. A QUAN-QUAL model was used, in which both qualitative and quantitative data was gathered and analyzed simultaneously and given equal weighting. This created a more holistic picture of the research findings (Roberts, 2010).

The quantitative aspect of the study consisted of a sample population of two 2nd grade classrooms (n=34) and two teachers. The teachers in the study collected the quantitative data over the course of the semester. The quantitative data was gathered in a pretest-posttest design. The data in the study consisted of the NWEA (Northwestern Evaluation Association) MAP reading RIT scores and percentile ratings, along with the Lexile ranges determined through the administration of the NWEA assessment. The NWEA MAP and Lexile ranges were collected from both the winter administration of the assessment and the spring administration. The pretest-posttest design allowed reading achievement to be compared from the beginning of the study until the end, as well as the growth that students made throughout this time period.

The teachers in the study not only collected the NWEA pretest-posttest MAP data and Lexile ranges, they also tracked other important pieces of data. They determined and collected the instructional reading levels (IRLs) of the students at the study at the beginning and end of the study with the Reading A-Z Benchmark books. The instructional reading levels that were found were then categorized to either be considered below, within, or above a 2nd grade reading level expectation, using the Reading A-Z chart. In addition, the teachers tracked how many weeks each student was instructed with a text for small group guided reading that was either below their IRL, that matched their IRL, or that was above their IRL. It is important to note that the teachers were not directed to use any particular level of text during small group guided reading, they merely tracked the decisions that they made throughout the study. Finally, gender, SES, Limited English Proficiency, race, and special education data was also collected for each student. These types of independent variables were analyzed alongside the NWEA MAP data in order to determine the possible connection to reading achievement and/or growth over the course of the study.

Descriptive and inferential statistics were run for the sample student population in order to determine any relationships between students' NWEA MAP and Lexile band growth or achievement from winter to spring, and the following: percentage of time students' spent reading text during guided reading that was below, that matched, and/or was above their instructional reading levels; whether or not students' instructional reading levels were categorized as below, on, or above the expected 2nd grade reading level range; gender; SES. For this study, race, Limited English Proficiency, and special education were not analyzed based on the school demographics. Chi-Square Linear-by-Linear Association and Likelihood Ratio Chi-Square were

used to establish any associations between the variables. Kruskal-Wallis was used to compare the mean ranks among the three groups of students based on the percentage of time they spent reading text during guided reading that was below, that matched, or was above students' instructional reading levels and their growth and achievement on NWEA MAP. When appropriate, the Mann-Whitney Pairwise Comparison post hoc test was used to determine where a statically difference may specifically exist between the three groups of students.

Following the initial brief observation of each classroom teacher, a series of three interviews were conducted with each teacher at the beginning, middle, and end of the study. Subsequently, the principal and the professional development liaison were interviewed twice, once at the beginning and again at the end of the study. First cycle coding, or open coding, was conducted in order to organize the data into common categories (Corbin & Strauss, 2007; Saldana, 2009). I used multiple methods when conducting first cycle coding; such as, descriptive coding, In Vivo coding, and structural coding (Saldana, 2009). Second cycle coding was then conducted in order to refine and reorganize the data codes from the first cycle into conceptual categories and themes. For the second cycle coding, I used axial coding to refine and further develop the codes that emerged from the first cycle coding process in order to determine broader themes (Corbin & Strauss, 2007). To enhance credibility and alleviate bias, my dissertation chair reviewed the coding throughout the processes of both first and second cycle coding.

Limitations of the Study

As in all research, mine has some limitations. However, this study did have limits geographically, demographically, and in sample size, which limited the ability for generalization. The data came from one school with a relatively homogeneous population. Geographically, the

study is from a Midwest state, and specifically one rural school district. Demographically, the student population at the district was predominately White with a moderate level of poverty. The sample size was relatively small; however, the mixed methods design of this study added richness to the data.

There were some limitations in this study based on its design. First of all, my original intent was to utilize an experimental design with randomly selected control and treatment groups, similar to the dyad study discussed in Chapter 2. However, I was unable to locate a school district that was comfortable with this type of research. Therefore, I adjusted my study to incorporate the professional judgment of teachers and included a method for them to track their decisions in order to determine what type of decisions had a more positive impact on reading achievement and growth. I also added a leadership component in order to help school leaders understand how to better support teachers when making literacy decisions.

Secondly, another design limitation emerged from the lack of alignment between the interview questions that were asked of the teachers and the school leaders throughout the study. When analyzing the data, it became challenging to find common categories between both groups, since there were several questions that were unique to each separate group. However, through the coding work that was completed some common categories and themes did emerge, which enabled me to do some qualitative correlation and synthesis.

Third, the fact that there were unequal amounts of students' initial reading level designations (whether they were considered reading within the second grade reading level, below the second grade reading level, or above the typical second grade reading level) at the beginning of the study was another limitation. If all the students had entered the study at the same level of

preparedness, or if there had been a more equal amount of students in each of the three categories, I would have been able to run inferential statistics in order to answer one of my research questions, which would determine if these designations had a statistically significant impact on their reading achievement or growth on NWEA.

The study had additional limitations with implementation. Inherently, since two different teachers participated in the study and they encompass varied levels of instructional effectiveness, training, and experience, this became a limitation. Even with the common training that was received through their work with the professional development liaison, each teacher's inherent bias impacts the practices of both teachers, including the amount and effectiveness of scaffolds provided to students. Secondly, since my study did not measure or collect information regarding the additional reading done by students at school or at home, I was unable to control for the type and frequency of reading students engaged in throughout the other times of day. Subsequently, the level of texts that were chosen to be read and the level of parental involvement with reading at home all impacted the results of this study and added to this limitation. Third, the parameters involving the length of the study was also a limitation. The assessment policy of the district set the parameters for the pre- and post- assessments. NWEA is given three times a year, once in August, second in January, and lastly in May. The approval necessary to find a school district in which to conduct the proposed research delayed its launch and prevented it from occurring over the entire school year. Therefore, the study occurred over one semester, which was about 18 weeks of instruction time.

Major Findings

The following section contains a succinct description of the findings supported in this research:

- Students who read texts during small group guided reading that matched their found instructional reading level between 47 - 67% of the time all increased their Lexile band from winter to spring within NWEA;
- Students who read texts during small group guided reading that were higher than their found instructional reading level up to 67% of the time met or exceeded the NWEA Spring RIT goal of 188;
- Teachers found that students were able to successfully access texts above their instructional reading levels when they provided scaffolds and supports;
- When determining text to use in guided reading, teachers shared the importance of considering texts that were interesting to students, along the instructional reading levels of the students and texts;
- Teachers found that when some students read text higher than their instructional reading level, they saw an increase in confidence as they tracked their own growth;
- Teachers became motivated to become more effective teachers when their students succeeded at something difficult or challenging, such as reading a text higher than their instructional reading level;
- School leaders and teachers identified the need to incorporate other assessments, in addition to NWEA, in order to better ascertain the instructional reading levels of students;

- Change in instructional practices seems to be best achieved when teachers actually see a positive impact on their students; and
- The school leaders shared that they believe they best impact the professional development of the teachers by assuming the role of facilitator.

Findings Related to the Literature

The findings from this mixed-methods study are presented in this section. A brief summary of the major findings is compared to the literature base. Some of the literature supports my findings and provides some evidence that students benefit from reading texts higher than their instructional reading levels. The findings are organized by research question, and the quantitative and qualitative data gathered were used to triangulate the results in order to drawn more rich and robust conclusions. When a finding impacted more than one research question, it was discussed alongside the first research question and then referenced within subsequent research questions.

Research Question 1. What is the relationship between the instructional reading levels of the texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading growth as measured by NWEA?

The first research question sought to determine if there was a statistically significant difference in reading growth within NWEA MAP reading RIT scores from winter to spring when compared to the percentage of time students read texts that were either below, that matched, or were higher than their instructional reading levels during guided reading. The NWEA Reading

MAP assessment served at the pre-post quantitative measure of student growth in reading throughout the study. After running multiple inferential tests, a statistically significant difference was not found between student growth on NWEA MAP Reading RIT scores from winter to spring and the amount of time students spent reading texts that were either below, that matched, or that were higher than their instructional reading levels.

However, when analyzing the Lexile band growth within NWEA from winter to spring against the percentage of time particular levels of instructional text were used with students, a statistically significant finding did appear (Table 14). This significant finding emerged between the percentages of time students read text during guided reading that matched their instructional reading levels and Lexile band growth. Once the Chi-Square displayed a statistically significant difference, and the Kruskal-Wallis approached significance, a Pairwise Comparison was completed. The post hoc Mann-Whitney indicated a statistically significant difference between the students in Group 3, who read texts that matched their IRL 47-67% of the time during the study, and the students in Group 2 who read texts that matched their IRL 20-33% of the time. The students in Group 3 all increased their Lexile band from winter to spring, and 66.7% of the students in Group 2 increased their Lexile band during this same time period. Therefore, this data analysis may indicate that students who spend more time reading texts during small group guided reading that match their instructional reading levels achieve more growth than those who read texts at their instructional reading level a lesser amount of time.

The qualitative aspect of this study provided additional meaning to this finding. Teachers shared their perceptions regarding their use of texts that matched students' instructional reading levels during small group guided reading. They felt as if it was important to have students read

some text during small group guided reading that was at students' instructional reading levels. Throughout the study teachers recorded that they had students read texts that matched their instructional reading levels on an average of 26% of the time. While this doesn't directly match the finding in this study that suggests students may want to read texts on their instructional reading level 47-57% of the time, it does indicate that the teachers chose to utilize texts that matched their students' instructional reading levels about a quarter of the time.

The conclusion that students make the greatest gains in small group guided reading when they spend time reading texts that match their instructional reading levels is supported by several findings in the literature review part of my study. As previously mentioned, having students read texts at their instructional reading levels in guided reading is and has been common practice for decades, originating from the work of Emmett Albert Betts (1946). Betts (1946) is considered the forefather of how to determine a student's independent, instructional, and frustration reading levels through his work on informal reading inventories, or IRIs. Through his work on IRIs he suggests that students who read texts that match their instructional reading levels will make greater gains in reading (Boley & Pennock 1975; Ekwall, 1976; Johns & Magliari, 1989; Pikulski, 1990; Powell & Dunkeld 1971; Williams, 1959). The conclusion that students achieve optimum growth from reading texts at their instructional reading level is also supported by most educators, reading specialists, and interventionists (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996), as well as some other researchers (Gickling & Armstrong, 1978; Treptow, Burns, & McComas, 2007). In conjunction with the above body of literature mentioned, the 2008 National Survey of Guided Reading Practices (Ford & Opitz, 2008) promoted the use of

instructional level reading materials as being the level of text to foster the most growth in reading achievement.

In summary, the finding that suggests that students make the greatest gains in reading growth when they spend time reading texts that match their instructional reading levels would not be considered an unexpected result. In fact, this finding is supported by a great deal of the literature on the subject (Boley & Pennock, 1975; Calkins, 2001; Clay, 2002; Ekwall, 1976; Fountas & Pinnell, 1996; Gickling & Armstrong, 1978; Johns & Magliari, 1989; Pikulski, 1990; Powell & Dunkeld 1971; Treptow, Burns, & McCormas, 2007; Williams, 1959) and backs the original research of Betts (1946), to some degree. The finding that emerged may indicate that in order for students to achieve the greatest gains in reading, students may want to read texts at their instructional reading levels for approximately half of the time during teacher-led small group guided reading, instead of the entire time.

Research Question 2. What is the relationship between the instructional reading level of texts used with students in small group guided reading that match, are higher, or are below the found instructional reading levels of the students and their end of semester reading achievement as measured by NWEA?

This research question was focused on whether or not a statistically significant relationship resided between the amount of time students spent reading texts in guided reading that were either below, that matched, or that were higher than their instructional reading levels and their achievement of a RIT score of 188 or higher on the spring NWEA MAP Reading assessment. The data analysis revealed that a statistically significant relationship did emerge between the percentages of time students spent reading texts higher than their instructional

reading levels and their achievement of the end of semester NWEA Spring RIT score of a 188 or higher. The Chi-Square and the Kruskal-Wallis inferential tests that were run both indicated a statistically significant relationship among the percentages of time students spent reading texts higher than their instructional reading levels and their achievement in NWEA. Furthermore, when the Pairwise Comparison was completed, there were two comparisons that showed significance. The students in both groups who read texts above their instructional reading levels up to 67% of the time displayed greater achievement on the NWEA spring RIT score, that the students who read text higher 73-100% of the time. This finding may indicate that if students read text higher than their instructional reading level up to 67% of the time during teacher-led small group guided reading they may gain greater achievement in reading.

The teachers in the study provided additional supporting data towards this finding during their interviews. As mentioned previously, the teachers collected data on the levels of texts used with students during small group guided reading throughout the study, and part of this included whether or the level of text used matched each students' instructional reading level, was above or was below. The average amount of time teachers chose texts for guided reading that were above students' instructional reading levels was about 60% of the time. In addition, an analysis of the data collected from both teachers showed that 90% of the students in the study read texts higher than their instructional reading levels for a minimum of two weeks. A qualitative finding that supports these quantitative results, is that the teachers in the study both shared that they saw benefits from using texts higher than students' instructional reading levels in small group guided reading, through an increase of growth and confidence. When discussing the use of texts above her students' instructional reading levels, she shared the following:

> But I do think, I've also seen for my ones that are growing a lot such a boost in confidence too. (For example)…when we sit down and do our running records and we are talking about, look, we read this level last time and now look where we are

This conclusion suggests that students of all levels of preparedness may need to read some texts higher than their instructional reading level to maximize their increase in reading achievement and is supported by some findings in my literature review. In their yearlong dyad (partner) reading study, Eldredge and Quinn (1988) found that when struggling second grade readers were given texts to read above their instructional reading level, described as a frustration reading level, with a different capable reader every week, they outperformed the control group who read texts at their instructional levels with partner reading. The findings of the study revealed that at the end of the experiment, 84% of the dyad reading experiment students scored on or above grade level on the Gates MacGinitie Reading Test; however, only 19% of the control group reached these levels of achievement (Eldredge & Quinn, 1988). In a subsequent study Morgan, Wilcox, and Eldredge (2000) conducted research in order to determine at what text difficulty level poor 2nd grade readers may make the greatest gains in dyad reading, and whether or not there is a point at which the level is too difficult even with the help of a partner. Their findings indicated that all three groups improved their reading ability; however, the students who read texts two levels above their found instructional reading levels made the most gains in reading growth, when compared to those who read at their level or who read four levels above their instructional reading level.

In conclusion, this result suggests that students who spend some time reading texts above their instructional reading levels during teacher-led small group guided reading may experience

greater reaching achievement. The supportive research in my literature review for this finding also indicated greater achievement in reading when more difficult texts were read by students in other literacy situations, such as dyad reading and fluency practice (Benjamin & Schwanenflugel, 2010; Eldredge & Quinn, 1988; Hiebert, 2005; Morgan, Wilcox, & Eldredge, 2000). It is important to note that these studies, including this result, are in conflict with the work of Betts (1946) and others (Boley & Pennock 1975; Calkins, 2001; Clay, 2002; Ekwall, 1976; Fountas & Pinnell, 1996; Gickling & Armstrong, 1978; Johns & Magliari, 1989; Pikulski, 1990; Powell & Dunkeld 1971; Treptow, Burns, & McCormas, 2007; Williams, 1959) who suggested that students should read texts within their instructional reading levels in order to make the greatest growth and achievement in reading.

Research Question 3. What is the relationship between whether students' found instructional reading levels are within, below, or above the typical grade level range and their end of semester reading achievement and growth on NWEA?

For this research question, I was unable to run inferential statistics in order to see if there was a statistically significant difference between whether a student entered the study with an instructional reading level considered below second grade reading level, within second grade reading level, or above second grade reading level (Calkins, 2001; Clay, 2002; Fountas & Pinnell, 1996). This was due to the uneven number of students designated within each of the three groups. At the beginning of this study, which was at the start of 2nd semester, approximately 76% of students were considered to be reading above second grade reading level, about 15% were considered to be reading within second grade reading level, and approximately 9% were considered to be reading below second grade reading level. Although not statistically

significant, according to the descriptive statistics calculated (Table 4; Table 8) students whose found instructional reading levels were above the typical 2nd grade reading level range attained more growth in their Lexile band (Benjamin, 2012) and met or exceeded the NWEA spring achievement score expectation at a much higher percentage than their peers, which is not an unexpected finding. In comparison, of the approximately 9% of students who entered the study reading below the 2nd grade reading level, texts higher than their instructional reading levels were read an average of 93% of the time during small group guided reading. Although not proven to be statistically significant, 67% of these below readers in the study were special education students. Furthermore, 67% of these below 2nd grade level readers exceeded both their projected growth in NWEA from winter to spring, and increased their Lexile bands within the same time frame.

In summary, even though this question was not able to be answered with statistically significant quantitative data as initially planned, the descriptive data may indicate that regardless of whether students entered the study reading above, at, or below the 2nd grade reading level expectation, they were able to access harder texts and experience reading growth. In order for me to have adequately answered this question, I would have needed a more robust sample size over a longer period of time in the field.

Research Question 4. What are the perceptions of the teachers as they increased the instructional reading level of texts used with some students?

Since an analysis of the quantitative data from the NWEA MAP Reading assessments from winter to spring would not provide data directly related to the perceptions of the teachers as they increased the instructional reading level of texts used with students during small group

guided reading, this question was intended to be answered using qualitative measures. This study was designed to track the professional judgement of the teachers regarding their decision making surrounding small group guided reading. As previously indicated, teachers both used the Reading A to Z Benchmarking system, along with information gained from NWEA, to determine the instructional reading levels of their students. Then, the teachers in the study used their professional judgment regarding the grouping of students and the levels of texts used for instruction. The teachers in the study both chose to use some texts higher than their students' instructional reading levels during guided reading. In fact, according to the data collected, 90% of the students in the study read texts above their instructional reading level for a minimum of two weeks.

One of the major findings that emerged was that both teachers found that students were able to access hard texts above their instructional reading levels during small group guided reading when they provided scaffolds and supports. The teachers both shared that when they utilized scaffolds, such as, graphic organizers, pre-teaching vocabulary, pre-reading, and building background knowledge through videos, the students were able to be successful with more difficult texts. One of the teachers summed up how she scaffolds differently for students so that they can successfully read texts at higher levels. She shared,

> And then for the groups that I pushed higher at a level, we just did a lot of background knowledge discussion, vocabulary and kind of reviewing our text features. (I used)…those strategies just to make sure they had front-loaded kind of what to expect. They'd also had a lot of exposure to life cycle and videos and stories and read-alouds and things. So …

> (they were) very familiar with the content, so that really helped them be successful as well I think.

The importance of an effective teacher's use of scaffolds in literacy instruction in order to help students be successful with complex texts is supported in research from my literature review. In their book, Text Complexity: Raising Rigor in Reading, Fisher, Frey, and Lapp (2012) stated, "We suggest that more difficult texts with scaffolded instruction should become part of the classroom equation. To ensure that students read complex texts, teachers have to revisit how they match readers with texts and tasks" (p. 5). Furthermore, effective teachers provide the needed scaffolds and instructional supports for students when instructing reading, which allows them to be more successful reading and accessing texts that may have initially been identified as a frustration level or hard (Fisher, Frey, & Lapp, 2012). Recent research has shown that an effective teacher has a high impact on student achievement (Marzano, Pickering, & Pollock, 2001; Reeves, 2010).

The use of scaffolds by the teachers in small group guided reading to support students in reading texts of increasing difficulty was found to be closely connected to an additional finding regarding teacher motivation. Through first cycle coding process of teacher interviews the category, *motivation to innovate GR strategies* emerged. The teachers shared that they were motivated to become more effective teachers when they observed their students succeeding at something difficult or challenging. One of the teachers in the study captured this thinking well when she shared the thoughts below.

> I think I see in my kids so much when they succeed, especially at something that's challenging or new to them, because that's kind of difficult for them. So when they actually

> do …(persevere) and they're so excited and so proud of themselves, not for doing something that's easy to them, but for doing something that's hard and that they overcame, I think that that pushes me …(because)I know that, OK, I'm responsible for them learning that.

The literature from Reeves, 2010 indicates that an effective teacher has a positive impact on student achievement. When you couple this with the literature from Fisher, Frey, and Lapp (2012) which suggests that an effective teacher provides the scaffolds for students in literacy instruction that may allow them to be more successful reading texts that are considered hard, it may suggest that teacher motivation is affected by student growth with difficult texts. This finding may support the use of higher level texts with teacher scaffolding, which in turn may increase the motivation, and effectiveness of teachers.

Finally, the teachers in the study both noted the added importance of utilizing student interest when determining which texts to have them read. This was a common theme both teachers. They both valued knowing what genres and subject matter the students were interested in, and they honored this whenever possible. One of the teachers shared the followed, which echoed the thoughts of both:

> I've tried to choose topics that they were interested in and it makes a difference when they can do some of the text selection. I can have them pick what text we do the next week out of a level basket, and that helps.

This is supported through the findings of Pitcher and Fang (2007) and Fisher, Frey, and Lapp (2012) who discussed that the importance of choosing a text is more than just semantics and syntax, and should also include the reader's age and interests, along with task and purpose.

Therefore, the more thoroughly teachers know the interests and backgrounds of their students, the more successful they will be when making decisions regarding which texts to match with which students (Pitcher & Fang, 2007). In addition, the historical link between reading achievement and student motivation, or interest, is highly important and is still considered an important aspect of reading instruction (Allington & Gabriel, 2012; Boushey & Moser, 2009; Ford & Opitz, 2008; Gambrell, 2011; Kennedy, 2010; Moley, Bandre, & George, 2011; . This finding may suggest that whenever possible, teachers should instruct students with more complex and interesting texts in order to increase reading engagement, which may in-turn have a greater impact on a student's reading achievement.

In conclusion, the perceptions that were noted among the teachers as they increased the difficulty of texts with students during small group guided reading were all supported by multiple sources found in the review of the literature. In summary, the teachers' use of scaffolds and supports with hard or difficult texts in order for students to be successful in reading them was reflected in the literature by Fisher, Frey and Lapp (2012). Secondly, the teachers' motivations to innovate within guided reading and the impact had on teacher effectiveness and student achievement is also supported by research (Fisher, Frey, & Lapp, 2012; Marzano, Pickering, & Pollock, 2010). Third, the importance that both teachers placed on student interest when choosing texts for guided reading has a large base of research for support in the link between students' interest and engagement in reading and their reading achievement (Allington & Gabriel, 2012; Boushey & Moser, 2009; Ford & Opitz, 2008; Gambrell, 2011; Kennedy, 2010; Moley, Bandre, & George, 2011; Pitcher & Fang, 2007).

Research Question 5. What are the perceptions of the principal and the professional development liaison in expanding the results of this study to other teachers?

This research question was designed to be answered by qualitative data since the analysis of the quantitative data from the NWEA MAP Reading assessments would not provide data directly related to the perceptions of the principal and the professional development liaison in expanding the results of this study to other teachers. Therefore, I relied on the qualitative data gathered and analyzed from the school leader interviews to answer Research Question 5. There were no major findings in my qualitative research that directly answered this question. However, a related major finding that the school leaders shared was that a need was uncovered through the participation in this study. The principal shared that through this study they realized that they needed to add another curricular tool or program to assist in identifying students' instructional reading levels and instructing small group guided reading. He shared,

> But moving forward, we recognized that some of the guided reading levels and the identification of those was a real concern, and so …(due) to the early literacy grant, we've now purchased some software that's going to aid all of our K-2 teachers in gathering that much-needed data. As you know, we were in year one of NWEA and we said this is a weakness of ours. So I think that's one thing that came out of this study to say we've got to have some additional tools.

There is literature to support the utilizing of more than just one assessment measure in order to determine instructional reading levels. Rubin (2011) recommended that teachers compare the scores from multiple assessments, such as running records and IRIs in conjunction with standardized assessments in order to make the most accurate decision on the instructional

reading levels of students. This thought was reiterated by Ford and Opitz (2008) when they wrote, "small group reading instruction of the past often relied solely on end-of-level assessments built into programs, but teachers implementing guided reading need to become increasingly more expert at continuously conducting assessments and interpreting results" (p. 322). They went on to encourage the use of a variety of assessment techniques when trying to ascertain a student's level of reading, honoring the complexity of the task.

This finding also supports the research of Guskey (2002) that true change in practice begins with professional development, which leads to a change of a teacher practice in the classroom, followed by a change in student learning outcomes, which results in a change in the attitudes and beliefs of the teachers. The teachers in this study received professional development, changed some instructional strategies in their practice, observed the impact on their students, and then adjusted their beliefs which, in this instance, resulted in the belief that they needed tools to ascertain instructional reading levels in addition to the NWEA data that they receive.

Additional finding not directly connected to a research question. Although not related to one of the research questions in this study, a related finding of importance was the perception of the school leaders' role in the professional development of the teachers. This was a finding that emerged from the analysis of the interviews of the school leaders in the study. The school leaders shared that they believe they best impact the professional development of the teachers by assuming the role of facilitator. For this study, the school leaders explained the role of facilitator as giving teachers permission to try something new, providing resources and training, being present in their classrooms, collaborating and working alongside teachers, and creating

opportunities to for them to share their learning with the staff. The principal shared the following,

> But then putting that in the hands of the teachers and letting them run with it. So, hey, here's a great idea. Here's a great new research tool. Let's explore this and let's run with it. I don't see my role in that process of dictating: you've …(got to) do this, you've …(got to) do that, you've …(got to)do this. But it's more of an opportunity to give them the tools, open their experiences to the new ideas and then let them run with them. Once we find it to be successful over the test market of certain classrooms, then it becomes more, hey, this is something we've got to do and this is the way we're going to put this into practice in our school. So it's still that role of a facilitator for me.

The research in my literature review supports this type of school leader role in professional development. The principal learning alongside teachers, being visible in the classroom, and creating a culture of reflection and continuous improvement, along with being knowledgeable about curriculum, instruction, and assessment, will maximize the learning of teachers, which in turn will maximize the learning of students (Fullan, 2014; Reeves, 2009;).

Finally, as an aside, the qualitative aspect of this study added to the richness of the research. The importance of the teachers' perspectives while instructing students at different levels of texts, along with the perception of the school leaders in the possibility of expanding the results of this study was very valuable. In addition, the school leaders ended up sharing their perceptions regarding their roles in the professional development of the teachers. This was not a specific research question, but something interesting that emerged. The qualitative portion included the two 2nd grade teachers in the study and the two school leaders, the principal and the

local university's professional development liaison. An observation of both teachers during their literacy block was conducted at the beginning of the study in order to better inform the research and gain an understanding of the context of small group literacy instruction in the classrooms, and not intended to answer a research question. As previously noted, these observations were able to validate the teachers' understanding and implementation of this study with some fidelity.

Surprises

There were some unexpected results that I noted as I completed my research. First of all, I was surprised when I realized that I would need to adjust my inferential statistical tests from parametric tests to nonparametric tests due to the sample size. The fact that I found two different statistically significant results regarding reading growth and achievement; one that indicated greater growth when using texts that matched students' instructional reading levels 47-67% of the time, and the other that indicated greater reading achievement when using text higher than students' instructional reading levels up to 67% of the time, was also unexpected. In addition, I was surprised when I realized that 90% of the students in the study were given text higher than their instructional reading level for a minimum of two weeks during the study. Furthermore, the fact that students spent an average of 60% of the time reading texts above their instructional reading levels, 26% of the time reading texts that matched their instructional reading levels, and an average of 8% of the time reading text below their instructional reading levels, was an additional unexpected finding. I thought the teachers would have chosen to have students spend the majority of the time reading text that matched their instructional reading levels during guided reading. The willingness of teachers to use text above students' IRLs to support this study was

surprising given the extensive prior research that indicates that optimal growth in reading is achieved through reading texts that match students' instructional reading levels.

Conclusions

The underlying question in this thesis was whether or not our current understanding of an instructional reading level is rigorous enough to maximize students' reading achievement and growth in the area of small group guided reading. This question was birthed from the difficulty in acquiring the original research used from Betts' (1946) *Foundations in Reading Instruction*, in which our past and current understanding of what accuracy percentages and comprehension levels constitute students' instructional reading levels. The results of this study appear to support the findings in the dyad reading studies in my literature review (Eldredge & Quinn, 1988; Morgan, Wilcox, & Eldredge, 2000) and Hiebert's (2005) fluency study that all indicate that students may need to read texts higher than their found instructional reading levels to maximize reading achievement and growth. Therefore, it is plausible that the answer to the question above is that our current understanding of what constitutes an instructional reading level may not be rigorous enough to maximize students' reading growth and achievement.

For decades educators have been afraid of moving kids above their instructional level and have avoided instructing students with what has been traditionally called hard or frustration level text for fear of stunting their reading growth (Betts, 1946; Ford & Opitz, 2008; Fountas & Pinnell, 1996; Gickling & Armstrong, 1978; Rubin, 2011; Treptow, Burns, & McComas, 2007). However, we may be preventing the maximum growth and achievement for students in reading by not giving them the opportunity to be stretched outside of their instructional range. With that being said, in order for the reading of frustration level texts to be successful, there are couple of

factors that emerged in my research that require consideration. First of all, teachers need to be active participants in guided reading in order to provide the scaffolds and supports necessary to students for them to successfully read and comprehend the more difficult text. This is supported through the research of Fisher, Frey, and Lapp (2012) stated, "We suggest that more difficult texts with scaffolded instruction should become part of the classroom equation. To ensure that students read complex texts, teachers have to revisit how they match readers with texts and tasks" (p. 5). Secondly, as supported by Fisher, Frey, and Lapp (2012) and Pitcher and Fang (2007), my research may also indicate that finding students' instructional reading levels is a starting point to determining which texts to use in small group instruction, but that student interests needs to be considered in order to fully impact reading achievement and growth. In order to increase a positive interaction between the reader and the text, teachers additionally need to know their students' interests and backgrounds in order to maximize reading growth and achievement (Donne, 2011; Fisher, Frey, & Lapp, 2012; Pitcher & Fang, 2007; Shanahan, 2010).

Finally, although not directly tied to this study, the finding that all students may achieve greater gains in reading when instructed with some texts above their instructional reading levels may need to be considered when striving to close the achievement gap (Martinez, Nellis, & Prendergast, 2006) of lower performing students. When considering the ability of the students in this study who were performing below grade level, they were able to successfully access harder levels of texts with the supports and scaffolds provided by their teachers. This study may suggest that even students who are struggling in reading, need to read texts both at and above their instructional reading levels with teacher support. We do not want struggling students to continually read texts below level because if that is the case, the gap between them and on-grade

level texts will continue to get larger. According to Fisher, Frey, and Lapp (2012) if a student is continually reading texts that are below his or her current grade level, then text difficulty is being decreased for that student over time, and he or she will not be able to access texts necessary at higher grade levels. Therefore, my research seems to suggest that a new understanding of the difficulty levels of texts used with all students during guided reading may need to be adjusted. In summation, for all students to maximize growth and achievement in reading it appears that they may need to read interesting texts that both match and are higher than their instructional reading levels, coupled with teacher scaffolds and supports.

Implications for Action

Since the analysis of my proposed study supported the dyad reading findings (Morgan et al., 2000), and oral reading findings (Hiebert, 2005) which indicated that utilizing texts above students' instructional reading levels may increase student achievement and growth, the following implications have resulted. First of all, the procedures to use when choosing texts for small group guided reading instruction may need to be adjusted. For example, my study suggests that students are able to read texts higher than their instructional reading levels when they are interested in the topic, and the teacher is providing scaffolds and supports throughout the process.

A connected implication may be the modification of teacher professional development regarding their practice of teaching small group guided reading. If further research confirms the findings in this study, teachers may require increased training in their ability to provide the necessary scaffolds and supports to students when reading texts above their found instructional reading levels. Furthermore, teachers would need altered professional development regarding the

structure of how often to utilize texts that match or are above a student's instructional reading level, while they are also considering student interest.

The findings of this study may also have implications on closing the achievement gap of lower performing students. This suggests that all students, even those reading below grade level, may be able to read text higher than their instructional reading levels if a teacher is providing the necessary scaffolds and supports. If further collaborated by further studies, this finding may alter teacher and practitioner practices in determining with texts to have students read in order to maximize their growth.

Recommendations for Future Research

One suggestion for further research would be for the study to be conducted with a larger sample size. A larger student sample, of at least 300 students, would potentially allow for more robust inferential statistics to be utilized, as explained in *Identifying and Implementing Educational Practice* by the U.S. Department of Education (2003). In turn, the increased sample size may allow for a greater opportunity to generalize the results to a larger population.

My original intent was to utilize an experimental design for my research with four second grade classrooms. In this design, half of the classrooms would have been randomly selected to be part of the control group, and the other half randomly selected to be in the experimental group. Based on the results of my study, I would recommend that the students in the experimental group be instructed with particular levels of texts both on and above their instructional reading levels for a specific amount of time. However, the control group would have read texts at their instructional reading levels, per teacher discretion throughout the study. The intent of this design

would be to more precisely determine the optimal amount of time students should read texts at and above a student's instructional reading level in order to yield the most growth. In addition, this study would strive to determine how many levels above a student's instructional reading level may influence the greatest gains. Therefore, I would recommend this type of experimental design be completed if there is a school corporation that would approve such a study.

Additionally, since my research may signify that current ranges of accuracy used to determine an instructional reading level may not be rigorous or extensive enough, I would recommend further research be done on the percentage ranges for reading accuracy used to determine the instructional reading levels of students. According to Powell (1971), very few investigations regarding the validity of the criteria used for determining the instructional reading level have been done. Also, Paolo (1977) shared that there is conflicting evidence regarding how the criteria for reading levels were established.

Finally, I would recommend extending a future study to expand across an entire school year. Developments in student growth were found over one semester in conjunction to reading certain levels of texts during small group guided reading and subsequent reading growth and achievement on NWEA. However, if the study had been able to be conducted over an entire school year, the trends may have been more apparent, and followed the typical time frame used by schools for one academic year.

Summary

In summation, this study considered the possibility that student reading growth may not be currently maximized based on the past and common understanding of what constitutes a

student's instructional reading level and the concurrent levels of text difficulty that are chosen to be used for instruction during teacher-led small group guided reading. Therefore, the main purpose of this study was to determine which instructional reading level text (one that matches, is below, or is above a student's found instructional reading level) used in small group guide reading may foster greater reading achievement and growth as measured by NWEA.

The results of this mixed methods study appear to indicate that an adjustment of our instructional practices surrounding the use of frustration level texts during teacher-supported small group guided reading may need to be reconsidered for all students, even those who are performing below grade level. Given my results, though limited in their ability to be generalized, I am cautiously optimistic that my study might be a catalyst for further research in this area. Subsequently, additional research may indicate that our instructional practices may need to be altered to incorporate the use of interesting texts that are of higher difficulty, currently considered frustration level, during guided reading with the addition of teacher scaffolds and supports in order to potentially help students maximize their growth in reading and close the achievement gap for struggling students.

References

Allington, R. L. & Gabriel, R. E. (2012). Every child, every day. *Educational Leadership, 69* (6), 10-15.

Allington, R. L., & McGill-Franzen, A. (2000). Looking back, looking forward: A conversation about teaching reading in the 21st century. *Reading Research Quarterly, 35*, 136-152.

Benjamin, R. (2012). Reconstructing readability: Recent developments and recommendations in the analysis of text difficulty. *Educational Psychology Review, 24*, 63-88.

Benjamin, R. G. & Schwanenflugel, P. J. (2010). Text complexity and oral reading prosody in young readers. *Reading Research Quarterly, 45* (4), 388-404.

Betts, E. A. (1946). *Foundations of reading instruction, with emphasis on differentiated guidance*. New York: American Book Company.

Boley, B. S. & Pennock, C. (1975). Crossfire. *The Reading Teacher, 28* (8), 782-783.

Boushey, G., Moser, J. (2009). *The café book: Engaging all students in daily literacy assessment and instruction.* Portland, ME: Stenhouse Publishers.

Calkins, L. M. (2001). *The art of teaching reading*. New York: Longman.

Clay, M. M. (2002). *An observation survey of early literacy achievement* (2nd ed.). Portsmouth, NH: Heinemann.

Cooper, J. L. (1952). *The effect of adjustment of basal reading materials on reading Achievement* (Unpublished doctoral dissertation). Boston University, Boston, MA.

Corbin, J., & Strauss, A. (2007). *The basics of qualitative research: techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: Sage.

Donne, V. (2011). Reading instruction and text difficulty. *The Volta Review, 111*, 5-23.

Dunn, M. (2010). Response to intervention and reading difficulties: A conceptual model that includes reading recovery. *Learning disabilities—A contemporary journal, 8*, 21-40.

Ekwall, E. E. (1976). Informal reading inventories: The instructional level. *The Reading Teacher, 29* (7), 662-665.

Eldredge, J. L., & Quinn, D. W. (1988). Increasing reading performance of low-achieving second graders with dyad reading groups. *Journal of Educational Research (82)*, 40-46.

Field, A. (2013). *Discovering statistics using ibm spss statistics* (4th ed.). Thousand Oaks, CA: Sage.

Fisher, D., Frey, N., & Lapp, D. (2012). *Text complexity: Raising rigor in reading.* Newark, DE: International Reading Association.

Ford, M. P. & Opitz, M. F. (2008). A National Survey of Guided Reading Practices: What we can learn from primary teachers. *Literacy Research and Instruction, 47* (4), 309-331.

Fountas, I. C. & Pinnell, G. S. (2009). *Benchmark assessment system.* Portsmouth, NH: Heinemann.

Fountas, I. C. & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

Fullan, M. (2014). *The principal: Three keys to maximizing impact*. San Francisco, CA: Jossey-Bass.

Gambrell, L. B. (2011). Seven rules of engagement: What's most important to know about motivation to read. *Reading Teacher, 65* (3), 172-178.

Gickling, E. E. & Armstrong, D. L. (1978). Levels of instructional difficulty as related to on-task behavior, task completion, and comprehension. *Journal of Learning Disabilities, 11* (9), 32-39.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Training: Theory and Practice, 8,* 381-391.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York, NY: Routledge.

Hiebert, E. H. (2011). Beyond single readability measures: Using multiple sources of information in establishing text complexity. *Journal of Education, 191* (2), 33-42.

Hiebert, E. H. (2005). The effects of text difficulty on second graders' fluency development. *Reading Psychology, 26* (2), 183-209.

Indiana Department of Education. (2016, September). *DOE compass.* Retrieved from http://compass.doe.in.gov/dashboard/overview.aspx

Johns, J. L., Magliari, A. M. (1989). Informal reading inventories: Are the betts criteria the best criteria? *Reading Improvement, 26* (2), 124-132.

Johns, J. L. (1991). Emmett a. betts on informal reading inventories. *Journal of Reading, 34* (6), 492-493.

Kennedy, E. (2010). Narrowing the achievement gap: Motivation, engagement, and self-

efficacy matter. *Journal of Education, 190* (3), 1-11.

Kontovourki, S. (2012). Reading leveled books in assessment-saturated classrooms: A close examination of unmarked processes of assessment. *Research Reading Quarterly, 47* (2), 153-171.

Learning A-Z Level Correlation Chart (2019). Retrieved February 24, 2019, from https://www.readinga-z.com/learninga-z-levels/level-correlation-chart/

Martínez, R. S., Nellis, L. M., & Prendergast, K. A. (2006). *Closing the achievement gap series: Part II,Response to Intervention (RTI) basic elements, practical applications, and policy recommendations.* (Education Policy Brief, Vol. 4, No. 8). Indiana University, Center for Evaluation and Education Policy

Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Moley, P. F., Bandre, P. E., & George, J. E. (2011). Moving beyond readability: Considering choice, motivation, and learner engagement. *Theory into Practice, 50* (3), 247-253.

Montana Office of Public Instruction. (October, *2016). Young scholars' academy for discovery and exploration balanced literacy checklist*. Retrieved from http://www.opi.mt.gov/streamer/profdev/13_10-18DWW_WestEd_Box/DWW%20Turning%20Around%20Chronically%20Low-

Performing%20Schools/Practice%20-%20Focus%20on%20Instruction/See/Sample%20Materials/2376_fi_mat_literacy.pdf

Morgan, A., Wilcox, B. R., & Eldredge, J. L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *Journal of Educational Research, 94* (2), 113-119.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards English Language Arts Appendix A.* Washington DC: National Governors Association Center for Best Practices & Council of Chief State School Officers.

Northwest Evaluation Association. (2016, July). *Map assessments.* Retrieved from https://www.nwea.org/assessments/map/

Northwest Evaluation Association. (2016, July). *2015 NWEA measures of academic progress normative data.* Available at: https://www.nwea.org/resources/2015-normative-data/

Paolo, M. F. (1977). *A comparison of readability graph scores and oral reading errors on trade books for beginning readers* (Unpublished doctoral dissertation). Rutgers University, New Brunswick, New Jersey.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Pikulski, J. J. (1974). A critical review: Informal reading inventories. *The Reading Teacher 28* (2), 141-151.

Pikulski, J. J. (1990). Informal reading inventories. *The Reading Teacher 43* (7), 514 - 516.

Pitcher, B. & Fang, Z. (2007). Can we trust leveled texts? An examination of their reliability and quality from linguistic perspective. *Literacy, 41*, 43-51.

Powell, W. R. & Dunkeld C. G. (1971). Validity of the IRI reading levels. *Elementary English, 48* (6), 637-642.

Pressley, M., Wharton-McDonald, R., Allington, R., Block, C. C., Morrow, L., Tracey, D., Baker, K., Brooks, G., Cronin, J., Nelson, E., & Woo, D. (2001). A study of effective first-grade literacy instruction. *Scientific Studies of Reading, 5*, 35-58.

Reeves, D. B. (2010). *Transforming professional development into student results.* Alexandria, VA: Association for Supervision and Curriculum Development.

Reeves, D. B. (2009). *Leading change in your school: How to conquer myths, build commitment, and get results*. Alexandria, VA: Association for Supervision and Curriculum Development.

Roberts, C. M. (2010). *The dissertation journey: A practical and comprehensive guide to planning, writing, and defending your dissertation* (2nd ed.). Thousand Oaks, CA: Corwin.

Robinson, V. M. J., Lloyd, C., & Rowe, K. J. (2008). The impact of educational leadership on student outcomes: An analysis of the differential effects of leadership types. *Education Administration Quarterly, 41*, 635-674.

Ross, J. A. (2004). Effects of running records assessment on early literacy achievement. *The Journal of Educational Research, 97* (4), 186-194.

Rubin, J. (2011). Organizing and evaluating results from multiple reading assessments. *Reading Teacher, 64* (8), 606-611.

Saldana, J. (2009). The coding manual for qualitative researchers. Los Angeles, CA: Sage.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed.). Boston, MA: Houghton Mifflin.

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: A practice guide* (NCEE 2010-4038). Washington DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from whatworks.ed.gov/publications/practiceguides.

Strong, J. H., Richard, H. B., & Catano, N. (2008). *Qualities of effective principals*. Alexandria, VA: Association for Supervision and Curriculum Development.

Thorndike, E.L. (1934). Improving the ability to read. *Teachers College Record, 36* (1), 1-19, (2), 123-144, (3), 229-241.

Treptow, M. A., Burns, M. K., & McComas, J. J. (2007). Reading at the frustration, instructional, and independent levels: The effects on students' reading comprehension and time on task. *School Psychology Review, 36*, 159-166.

USA Factbook. (2016, September). Retrieved from: http://globalsearchusrs.ca.com/cities/indiana/cowan

U.S. Department of Education: Institute of Education Sciences. (2003). *Identifying and implementing educational practices supported by rigorous evidence: a user friendly guide*. Retrieved from https://www2.ed.gov/rschstat/research/pubs/rigorousevid/rigorousevid.pdf

Wigfield, A., Guthrie, J. T. (1997). Motivation for reading: Individual, home, textual, and classroom perspective. *Educational Psychologist, 32*, 57-135.

Williams, G. (1959). Provisions for critical-reading in basic readers. *Elementary English, 36* (5), 323-331.

Appendix A NWEA MAP for Primary Grades Conversion Table

2019 Learning A–Z Correlation Chart

Learning A–Z	Grade	Ages	Lexile®	Accelerated Reader (ATOS)	DRA	Fountas & Pinnell	Reading Recovery	PM Readers
■	K	4–6	BR70L–10L	0–.9	A–1	A	1	Starters 1
A	K	4–6		0–.9	A–1	A	1	Starters 1
B	K	4–6	BR40L–160L	0–.9	2	B	2	Starters 2
C	K	4–6		0–.9	3–4	C	3–4	3–4 red
D	1	4–7	160L–310L	1–2.4	6	D	5–6	5–6 red/yellow
E	1	6–7		1–2.4	8	E	7–8	7–8 yellow
F	1	6–7	300L–450L	1–2.4	10	F	9–10	9–10 blue
G	1	6–7		1–2.4	12	G	11–12	11–12 blue/green
H	1	6–7	430L–530L	1–2.4	14	H	13–14	13–14 green
I	1	6–7		1–2.4	16	I	15–16	15–16 orange
J	1	6–8		1–2.4	18	J	17	17 turquoise
K	2	7–8	510L–620L	2.5–3.5	18	J	17	18 turquoise
L	2	7–8		2.5–3.5	20	K	18	19–20 purple
M	2	7–8	530L–810L	2.5–3.5	24	L	19	21 gold
N	2	7–8		2.5–3.5	28	M	20	22 gold
O	2	7–8	600L–850L	2.5–3.5	28	M	20	22 gold
P	2	7–8		2.5–3.5	28	M	28	22 gold
Q	3	7–9	660L–930L	3.6–4.2	30	N	30	23 silver
R	3	8–9		3.6–4.2	30	N	30	23 silver
S	3	8–9	790L–940L	3.6–4.2	34	O	34	24 silver
T	3	8–9		3.6–4.2	38	P	38	25 emerald
U	4	8–11	820L–1030L	4.3–4.9	40	Q	40	26 emerald
V	4	9–11		4.3–4.9	40	R	40	26 emerald
W	4	9–11		4.3–4.9	40	S	40	27 ruby
X	5	9–11	890L–1080L	5.0–5.5	40	S	40	28 sapphire
Y	5	9–11		5.0–5.5	40	T	40	29 sapphire
Z	5	9–11		5.0–5.5	50	U–V	N/A	30 sapphire
Z1	5+	9–11+	920L–1120L	5.6–6.3	60	W–X	N/A	N/A
Z2	5+	9–11+		6.4–6.9	70+	Y–Z	N/A	N/A

This correlation chart illustrates how Learning A–Z levels approximately correlate to other leveling systems commonly found in leveled reading materials. Learning A–Z uses objective (quantitative) and subjective (qualitative) Leveling Criteria to measure text complexity. Use this chart to identify correlated levels for Raz-Plus, Reading A–Z, and Raz-Kids resources. ®Lexile® bands are certified through a partnership with MetaMetrics®.

Learning A–Z

Appendix B NWEA MAP for Primary Grades Conversion Table

NWEA
Northwest Evaluation Association
Partnering to help all kids learn

NWEA MAP for Primary Grades Conversion Table: Fountas and Pinnell

NWEA used a technique called an "Equipercentile Method" to equate MPG scores to levels on Fountas & Pinnell which have different scales. The method has been found to accurately predict the MAP RIT scale to other instruments (Cronin, et al., 2007). Simply stated, the Equipercentile method looked at the distributions of scores for students who took both the MPG test and were scored on Fountas & Pinnell. This provides a MAP RIT score or range of scores which are equivalent to the F & P levels. For more information on this method, please see the research by Ryan and Brockmann, 2009.

NWEA MAP for Primary Grades RIT Range	Fountas & Pinnell Reading Level
90 - 141	PC
142 - 147	RB/A
148 - 155	B
156 - 161	C/D
162 - 166	E/F
167 - 173	G/H
174 - 179	I/J
180 - 185	K
186 - 191	L
192 - 196	M
197 - 199	N
200 - 201	O
202 - 350	P

Northwest Evaluation Association (NWEA) is a not-for-profit corporation with a mission to help all kids learn. This conversion table is provided as a reference tool only to help teachers, students and parents recommend books to encourage reading. NWEA has not verified the correlation of the reading levels to NWEA's RIT score and is providing the conversion table "as is" without warranty.

The names, products and services of other companies are the property of their respective owners. The F&P Text Level Gradient™ is a product of Fountas & Pinnell.

Appendix C Small Group Guided Reading Initial Observation Rubric

Is there evidence of purposeful planning for reading groups based on current data?	Not Met	On the way	Met	Notes
Evidence of planning				
Teacher has access to a variety of leveled texts and genre from a school source				
Scholars are grouped based on reading assessment				
The text selected is appropriate for small group instruction				
The text is usually an unfamiliar text				
The lesson has a focus that includes the strategy to be practiced				
Introduction of the text activates prior knowledge and discussion on the topic including unusual vocabulary				
Thinking, talking, questioning through text				
Purpose is set for each segment of reading				
Scholars reading independently (not round robin)				
Discussion after reading of each section				
Teacher observing scholars' reading behaviors and takes conference notes				
Meaningful extension activities provided for early finishers related to the text				
A reading strategy given and pursued by scholars				
Scholars are given the opportunity to reread the text				
Pacing of lesson is appropriate				
A variety of texts used over time and evident				
Connections made to how strategy can be used in independent reading, and is observed in subsequent independent reading sessions				
A follow up literacy activity can be given				

Appendix D Teacher Interview Protocol Form

Date: _____

Interviewer: Stacy Smith

Interviewee (Title and Name): _____

Post Interview Comments or Leads: _____

Introductory Protocol

Thank you very much for agreeing to meet with me. In order to be able to use active listening and also capture all of the details, I would like to audio tape our conversation. After the interview is transcribed, the audio tape will be deleted. In addition, you must sign a form devised to meet the human subject requirements. Essentially, this document states that: (1) all information will be held confidential, (2) your participation is voluntary and you may stop at any time if you feel uncomfortable, and (3) I do not intend to inflict any harm. Once again, thank you very much for being willing to participate.

I value your time and therefore, I have planned for this interview to not last longer than one hour. I do have several questions that I would like to ask you during this time. If for some reason, our time is running out, I might need to interrupt you in order to stay within the hour and complete the line of questioning.

Introduction

You have been selected to be a part of this interview since you are one of the teachers involved in my research project regarding the relationship between the instructional reading levels used with students in small group guided reading and their reading achievement and growth. More specifically, I am curious if there is a difference in reading achievement and growth when using texts that match a student's instructional reading level verses using a text above a student's instructional reading level in small group guided reading.

Core Questions for All Three Interviews

1. How are you feeling about your small group guided reading instruction?

2. What tools do you use to determine the instructional reading levels of your students?

   Prompt: Are the tools you have used still effective? What changes have you made? (middle and final interviews)

3. Describe the tools or observations you rely on when deciding how to match small groups of students with levels of texts used in small group guided reading.

   Prompt: What changes have you made since? (middle and final interviews)

4. Tell me about the decision making process you go through when deciding how to group students for small group guided reading.

   Prompt: How has your decision-making changed? (middle and final interviews)

5. What types of scaffolding do you use when instructing the students that are reading a text above their instructional reading level?

   Prompt: How were the scaffolds the same or different from those used with students reading texts at their instructional reading level?

6. What might be some strategies you have used to help students access the texts they are reading?

   Prompt: What did you notice about the effectiveness of those strategies?

7. What differences emerged in student behaviors in those who were reading text on-level, above-level, or even below-level?

   Prompt: What are your hunches about what caused those differences?

8. Describe any difficulties that were encountered.

   Prompt: Benefits?

9. As you reflect on your experiences, what did you pay attention to as you differentiated?

   Prompt: What decisions did you make about how to monitor and adjust?

10. What are some ways your interactions with your colleagues have assisted you in your guided reading instruction?

    Prompt: How have your interactions with your school leaders assisted you? (Principal and professional development liaison?

Additional Question for Initial Interview:

1. What motivates you to incorporate innovative instructional strategies in small group guided reading? (initial)

Additional Question for Mid-Study Interview:

1. A follow-up question regarding the initial teacher observations of small group guided reading will be asked in order to clarify literacy practices.

Additional Questions for Final Interview:

2. What suggestions might you have for the future implementation of this study?

3. What do you want to stay mindful of from now on regarding your guided reading practices?

4. How might you ensure that you maintain focus on these areas?

Ending Protocol

Thank you very much for your time and sharing your thoughts with me. I might need to contact you again in order to clarify information or ask some additional questions. You can reach me at ssmith4@ccs.k12.in.us or 317-385-8491.

Appendix E School Leader Interview Protocol Form

Date: _____

Interviewer: Stacy Smith

Interviewee (Title and Name): _____

Post Interview Comments or Leads: _____

Introductory Protocol

Thank you very much for agreeing to meet with me. In order to be able to use active listening and also capture all of the details, I would like to audio tape our conversation. After the interview is transcribed, the audio tape will be deleted. In addition, you must sign a form devised to meet the human subject requirements. Essentially, this document states that: (1) all information will be held confidential, (2) your participation is voluntary and you may stop at any time if you feel uncomfortable, and (3) I do not intend to inflict any harm. Once again, thank you very much for being willing to participate.

I value your time and therefore, I have planned for this interview to not last longer than one hour. I do have several questions that I would like to ask you during this time. If for some reason, our time is running out, I might need to interrupt you in order to stay within the hour and complete the line of questioning.

Introduction

You have been selected to be a part of this interview since you are the principal/professional development liaison at the building were my research project regarding the relationship between the instructional reading levels used with students in small group guided reading and their reading achievement and growth is occurring. More specifically, I am curious if there is a difference in reading achievement and growth when using texts that match a student's instructional reading level verses using a text above a student's instructional reading level in small group guided reading.

1. Briefly describe why you chose education as your profession?

2. Tell me about the structures put in place in order to foster teacher-leader collaboration and reflection regarding student achievement data.

3. Describe your role as an instructional leader in the school?

   a. Probe: Tell me about your on-going professional development, particularly in literacy instruction.

4. What motivates you to encourage innovation in instruction with your teachers?

5. What additional tools or materials did you provide for the teachers who were a part of the study and why?

6. Tell me how you can best support teachers as they implement literacy innovations?

7. What role has your visibility in the classroom and formative feedback given to teachers played in this study?

Additional Questions for Final Interview: (Share preliminary results)

8. Describe your perceptions regarding expanding the results of this study to other teachers?

Ending Protocol

Thank you very much for your time and sharing your thoughts with me. I might need to contact you again in order to clarify information or ask some additional questions. You can reach me at ssmith4@ccs.k12.in.us or 317-385-8491.